Title
Regulating TNCs: Should Uber and Lyft set their own rules?

Authors
Li, S
Tavafoghi, H
Poolla, K
et al.

Publication Date
2019-11-01

DOI
10.1016/j.trb.2019.09.008

Peer reviewed

Regulating TNCs: Should Uber and Lyft Set Their Own Rules?

Sen Li[a], Hamidreza Tavafoghi[a], Kameshwar Poolla[a,b], Pravin Varaiya[b]

[a] *Department of Mechanical Engineering, University of California, Berkeley*

[b] *Department of Electrical Engineering and Computer Science, University of California, Berkeley*

Abstract

We evaluate the impact of three proposed regulations of transportation network companies (TNCs) like Uber, Lyft and Didi: (1) a minimum wage for drivers, (2) a cap on the number of drivers or vehicles, and (3) a per-trip congestion tax. The impact is assessed using a queuing theoretic equilibrium model which incorporates the stochastic dynamics of the app-based ride-hailing matching platform, the ride prices and driver wages established by the platform, and the incentives of passengers and drivers. We show that a floor placed under driver earnings pushes the ride-hailing platform to hire *more* drivers and offer *more* rides, at the same time that passengers enjoy *faster* rides and *lower* total cost, while platform rents are reduced. Contrary to standard economic theory, enforcing a minimum wage for drivers benefits both drivers and passengers, and promotes the efficiency of the entire system. This surprising outcome holds for almost all model parameters, and it occurs because the wage floors curbs TNC labor market power. In contrast to a wage floor, imposing a cap on the number of vehicles hurts drivers, because the platform reaps all the benefits of limiting supply. The congestion tax has the expected impact: fares increase, wages and platform revenue decrease. We also construct variants of the model to briefly discuss platform subsidy, platform competition, and autonomous vehicles.

Keywords: TNC, wage floor, ride-haling tax, regulatory policy.

1. Introduction

In December 2018, New York became the first US city to adopt a minimum wage for drivers working for app-based transportation network companies (TNCs) like Uber and Lyft. The New York City Taxi and Limousine Commission (NYTLC) established a "minimum per-trip payment formula" that gives an estimated gross hourly driver earnings before expenses of at least \$27.86 per hour and a net income of \$17.22 per hour after expenses, equivalent to the minimum wage of \$15 per hour because, as "independent contractors," drivers pay additional payroll taxes and get no paid time off [1]. The NYTLC formula for non-wheelchair accessible vehicles is

$$\text{Driver pay per trip} = \left(\frac{\$0.631 \times \text{Trip Miles}}{\text{Company Utilization Rate}}\right) + \left(\frac{\$0.287 \times \text{Trip Minutes}}{\text{Company Utilization Rate}}\right) + \text{Shared Ride Bonus} \tag{1}$$

Email addresses: `lisen1990@berkeley.edu` (Sen Li), `tavaf@berkeley.edu` (Hamidreza Tavafoghi), `poolla@berkeley.edu` (Kameshwar Poolla), `varaiya@berkeley.edu` (Pravin Varaiya)

amounting to $23 for a 30-min, 7.5-mile ride.[1] New York City's $15/hour minimum wage for large employers, which went into effect on December 31, 2018 doesn't apply to drivers who work for ride-hailing apps.

The Commission imposed this wage floor based on testimony on driver expenses, meetings with stakeholders, and on the report of labor economists J.A. Parrott and M. Reich [2] which showed that median driver earnings had declined almost $3.00 per hour from $25.78 in September 2016 to $ 22.90 in October 2017, a decrease of 11.17%. The TNCs imposed the $3.00 per hour wage cut during a period when the number of drivers in the largest four TNCs (Uber, Lyft, Gett/Juno, and Via) had grown by 80,000 [1]. Uber would be the largest for-profit private employer in New York City if its drivers were classified as employees rather than independent contractors [2]. The ingenious wage formula (1) encourages TNCs to increase driver pay through higher utilization, instead of trying to restrict the number of drivers through regulation. Lyft, however, opposed the regulation saying that because of its larger size, Uber's higher utilization rate gave it an unfair advantage [3]. Lyft's complaint was overruled [4].

The subminimum wage of drivers working for TNCs also prompted the Seattle City Council in April 2018 to pass a unanimous resolution to explore setting a minimum base rate of $2.40 per mile for TNCs compared with the prevailing rate of $1.35 per mile and the rate of $2.70 per mile charged by taxis [5]. The resolution also asked TNCs to voluntarily hand over anonymous data on hours, trips, fares and compensation. Unlike NYTLC, however, no other US city has access to TNC data to estimate what their drivers are paid or the TNC impact on traffic. For example, the California Public Utilities Commission which regulates TNCs will not share TNC data with San Francisco County Transportation Authority [6]. TNC regulation "follows an elite political process dominated by concentrated actors and government decision makers largely acting ex officio (committee heads, regulators, and judges)" [7].

In December 2018, Uber lost its case at the U.K. Court of Appeal against the October 2016 ruling that its drivers should be classified as workers entitled to rights such as minimum wage and paid holidays. The Court ruled against Uber's claim that its drivers were just self-employed contractors who use its app in exchange for a share of their fares at the level dictated by Uber [8]. The case can be used to challenge the self-employed status of millions of gig-economy workers who work for companies like Airbnb and Deliveroo on a freelance basis without fixed contracts. New York and London are the largest Uber markets in the US and EU. The California state assembly recently passed bill AB5 that would make hundreds of thousands of independent contractors including TNC drivers become employees. The bill now goes to the senate [9]. Uber and Lyft are aggressively campaigning against AB5. In its SEC filing, Uber states "If, as a result of legislation or judicial decisions, we are required to classify Drivers as employees . . . we would incur significant additional expenses [that would] require us to fundamentally change our business model, and consequently have an adverse effect on our business and financial condition [10, p.28]."[2]

As of January 1, 2019, all trips by *for-hire* vehicles that cross 96th street in NYC will pay a congestion surcharge of $2.75 per TNC trip, $2.50 per taxi trip, and $0.75 per pool trip. Further, NYC will also charge a toll on *every* vehicle that enters the busiest areas, currently defined as south of 61st street. This 'cordon' price will raise about $1B per year (assuming a $11.52 toll) for the Metropolitan Transportation Authority.

Uber's reaction to these adverse decisions was predictable. Responding to the NYTLC ruling Uber's director of public affairs stated, "legislation to increase driver earnings will lead to higher than necessary fare increases for riders while missing an opportunity to deal with congestion in Manhattan's central business district" [12].[3] Uber challenged the Seattle resolution: its general manager for Seattle said, "we are generally unclear

[1]The utilization rate is calculated as the total amount of time drivers spend transporting passengers on trips dispatched by the base divided by the total amount of time drivers are available to accept dispatches from the base [1]. Wheelchair accessible vehicles receive a higher rate.

[2]For a thoughtful discussion of labor-market trends in the gig-economy see [11].

[3]Lyft echoed the Uber response stating, "These rules would be a step backward for New Yorkers, and we urge the TLC to

how nearly doubling per-mile rider rates would not result in an increased cost for riders"[5]. Uber also declared it would fight the U.K. Appeal Court's decision in the Supreme Court [8]. Contradicting Uber's claims, this study shows that raising driver wages will *increase* the number of drivers and riders at the same time that passengers enjoy *faster* rides and *lower* total cost, while platform rents are reduced.

The aforementioned regulations are part of the political response to the public anxiety over the disruption of the urban transportation system caused by the rapid growth of TNCs. Worldwide, the monthly number of Uber users is forecast to reach 100 million in 2018, up from 75 million in 2017. In New York, the four largest TNCs Uber, Lyft, Juno and Via combined dispatched nearly 600,000 rides per day in the first quarter of 2018, increasing their annual trip totals by over 100 percent in 2016 and by 71 percent in 2017. About 80,000 vehicles are affiliated with these four companies [2]. In San Francisco, 5,700 TNC vehicles operate in peak times. They daily make over 170,000 vehicle trips, approximately 12 times the number of taxi trips, and 15 percent of all intra-San Francisco trips, comprising at least 9 percent of all San Francisco person trips [13]. This explosive growth of TNCs has raised two public concerns.

As noted earlier, one concern is with the working conditions of TNC drivers. The TNC business model places much of the economic risk associated with the app sector on drivers, who are classified as independent contractors. Furthermore, the model relies on having many idle cars and drivers, resulting in low driver pay per hour and high TNC platform rents.[4] TNCs need idle drivers to reduce passenger waiting time. Uber's annual revenue from passenger fares in New York City amounts to about $2 billion, of which it keeps about $375 million in commissions and fees, for a markup estimated at six times its variable operating cost or 600 percent [2]. One common opinion is that "Uber's driver-partners are attracted to the flexible schedules that driving on the Uber platform affords . . . because the nature of the work, the flexibility, and the compensation appeals to them compared with other available options [14]." In fact, more than 60 percent of New York City drivers work full-time and provide 80 percent of all rides; their work hours are not flexible [2].

The second concern is with the negative impact of TNCs on a city's traffic congestion and its public transit ridership. A detailed 2017 report [15] examined the impact of TNC growth on traffic conditions in Manhattan's CBD. The analysis shows that, from 2013 to 2017, TNC trips increased 15 percent, VMT increased 36 percent, traffic speed declined 15 percent, the number of vehicles increased 59 percent, and the number of unoccupied vehicles increased 81 percent. The report suggested reducing the unoccupied time of TNC vehicles as a means of congestion control. Responding to the increased congestion, the New York City Council in 2018 passed a regulation freezing the number of TNC vehicles on the road for one year. Supporters of the cap, including Mayor Bill de Blasio, said the regulation will protect drivers, fairly regulate the industry and reduce congestion [16]. However, our analysis shows that imposing a cap hurts drivers, because the TNC retains as profit the benefits of limiting supply.

Another detailed report [13] by San Francisco Transportation Authority provides information on the size, location, and time-of-day characteristics of TNC activities in San Francisco. A follow-up report [17] identifies the impact of TNC activities on road congestion in San Francisco. It shows that after subtracting the impact of employment growth, population change and network capacity change, TNCs contributed 51 percent of the increase in vehicle hours of delay, 47 percent of increase in VMT, and 55 percent of the average speed decline between 2010 and 2016. Moreover, "TNC trips are concentrated in the densest and most congested parts of San Francisco including the downtown and northeastern core of the city. At peak periods, TNCs are estimated to comprise 25 percent of vehicle trips in South of Market." The report cites studies showing that "between 43 percent and 61 percent of TNC trips substitute for transit, walk, or bike travel or would not have been made at all."

---

reconsider them [12]."

[4] TNC expenditures comprise a fixed initial cost for setting up the platform and a small variable cost as the company grows. Thus the average cost per trip falls and its profit margin increases as the TNC grows.

This paper evaluates three TNC regulations: a minimum driver wage, a cap on the number of drivers or vehicles, and a per-trip congestion tax. We analyze the impacts of these regulations on several aspects of the app-based ride-hailing market, including ride prices and driver wages established by the platform, the incentives of passengers and drivers, vehicle occupancy, and platform rent or profit. We use a model to determine the arrival of passengers, number of drivers, ride prices and platform commissions, conditioned on the imposed regulation. The model employs a queuing theoretic model with dynamic matching of passengers and drivers, an equilibrium model that predicts the long-term average arrivals of passengers and drivers, and an optimization model of platform decision-making. We summarize the key results.

- Imposing a minimum wage will motivate TNCs to hire *more* drivers and offer *more* rides, and passengers to enjoy *faster* rides and *lower* total cost, while TNC rent or profit shrinks. It indicates that raising the minimum wage will benefit both drivers and passengers, while platform rent will decline. This counter-intuitive result holds for almost all model parameters, and it occurs because the wage floor curbs TNC labor market power.
- Contrary to common belief, a cap on the number of drivers will hurt driver earnings. This is because when fewer drivers are permitted, the platform will hire cheaper labor by reducing driver pay. Thus, the benefit of limiting the driver supply is retained by the platform.
- Imposing a congestion surcharge has a predictable impact: the numbers of passengers and drivers and the platform revenue reduce as the congestion surcharge increases. Our numerical study shows that a congestion surcharge of \$2.75/trip significantly reduces the platform profit in NYC. This suggests that the business model of TNC is vulnerable to the adverse effect of congestion policies.

We also present variants of our model to analyze platform subsidy, platform competition and autonomous vehicles.

Related Work: There are several studies of ride-hailing platforms. A recurrent concern is to evaluate decisions that maximize platform profit, with particular attention to static vs. dynamic pricing. A queuing model is proposed in [18] to study the profit maximizing prices of ride-hailing platforms. It shows that the throughput and profit under dynamic pricing strategy can not exceed that under the optimal static pricing strategy that is agnostic to stochastic dynamics of demands. On the other hand, dynamic pricing is more robust to fluctuations in system parameters compared to static pricing. Hence, the platform can use dynamic pricing to realize the benefits of optimal static pricing without perfect knowledge of system parameters.

A similar question is studied in [19], with a focus on the self-scheduling capacity of for-hire drivers. It is shown that the additional flexibility of drivers is beneficial to platforms, consumers and drivers. It also suggests that when some periods have predictably higher demand than others (e.g., a rainy evening), with static pricing it is hard to find service at peak demand times, so surge pricing is likely to benefit all stakeholders. In the same vein, [20] suggests dynamic pricing for the platform to maximize the profit across different time periods when the underlying operating characteristics change significantly. It is shown in [21] that platform pricing can be more complicated when there is uncertainty in passenger's valuation or driver's opportunity cost. A general economic equilibrium model is developed in [22] to evaluate the impacts of ride-hailing services on deadhead miles and traffic congestion. Ride-hailing platforms are also examined as a special kind of two-sided platforms. See [23] and [24] for a summary of literature on two-sided platforms, and [25] for a general theory of monopoly pricing in multi-sided platforms.

The literature on regulation of the app-based ride-hailing marketplace is relatively limited. A ride-hailing platform that manages a group of self-scheduling drivers to serve time-varying demand is studied in [26]. The study shows that under a wage floor, the platform starts to limit agent flexibility because it limits the number of agents that can work in some time intervals.

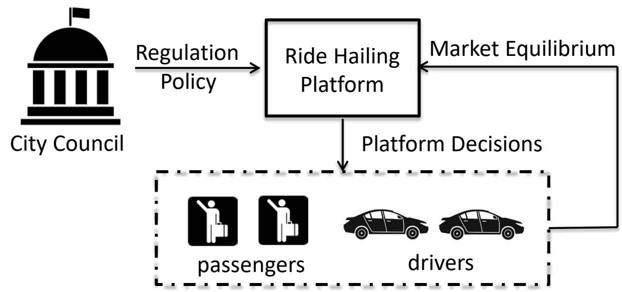

Figure 1: The TNC system includes the city council, platform, passengers and drivers.

The work closest to ours is by Parrott and Reich [2]. The authors use TNC administrative data collected by the New York City Taxi and Limousine Commission (NYTLC) to examine the likely impact of the NYTLC's proposed regulations [1]. By numerical simulation, they show that the proposed policy will increase driver earnings by 22.5 percent, while passengers will only experience moderate increase of trip fare (less than 5 percent) and waiting times (12 to 15 seconds). However, our analysis shows that both the trip cost and the waiting time will decrease. This is because in our model we assume that the passengers are sensitive to the pickup time of the ride-hailing services, which is not captured in [2].

2. TNC Environment

This section describes the TNC enviroment. Agents of the transportation system are comprised of the city council, the app-based ride-hailing platform (TNC), a group of passengers and for-hire vehicle drivers (see Figure 1). The city council sets legislation to regulate TNC operations. Examples of regulations include minimum driver wage, maximum number of vehicles and regional licensing.[5] The regulations are enforced by auditing the operational data of TNCs. (See [1] for details of enforcement in New York.) The platform responds to the regulations by setting profit-maximizing ride fares and driver commissions (or equivalently, wages). These fares and wages are called 'platform decisions' in Figure 1. The platform decisions influence the choices of passengers and drivers. For instance, passengers have diverse ride choices including TNC, public transit, walking, and biking. They select an option based on the cost and convenience of each choice. Drivers also have alternative job opportunities, such as delivering food, grocery, packages, and mail. They take the job with the highest expected wage. The choices of passengers and drivers form a market equilibrium, which determines the TNC profit or rent. The equilibrium is affected by regulations.

The objective of the paper is to understand how regulations impact the ride-hailing transportation system. We consider three regulations: (a) a floor under driver wage; (b) a cap on total number of drivers; and (c) a per-trip congestion tax. We analyze their impact from various perspectives of the ride-hailing system, including ride fares, commission rate, passenger pickup time, driver earnings, platform rent, number of riders, number of for-hire vehicles, and vehicle occupancy rate.

The rest of this paper is organized as follows. In Section 3 we introduce the market equilibrium model of the response of passengers and drivers to a platform decision. In Section 4 we predict TNC decisions in the absence of regulation. In Section 5 we examine TNC decisions operating with a floor under driver wage rate. In Section 6 we consider TNC decisions when there is a cap on the number of drivers.In Section 7 we study the impact of congestion surchage. Platform competition and other model variations are discussed in Section 8. Conclusions are offered in Section 9. Several proofs are deferred to the appendix.

[5]Unlike TNCs, taxicabs are heavily regulated.

3. Market Equilibrium Model

We now develop the market equilibrium model of the decisions of drivers and passengers in response to the platform decision. The model is used to predict the average arrival rates of passengers and number of drivers.

3.1. Matching Passengers and Drivers: M/G/N Queue

We use a continuous-time queuing process to model the matching of passengers and drivers. Consider N TNC drivers or vehicles, each modeled as a server. A vehicle is 'busy' if there is a passenger on board, or a passenger is assigned and the vehicle is on its way to picking her up. Otherwise, it is considered 'idle'. We assume that the arrival process of passengers is Poisson with rate $\lambda > 0$. Newly arrived passengers immediately join the queue and wait until an idle vehicle is dispatched by the platform. Hence this is an M/G/N queue, and the expected number of idle servers (vehicles) is $N_I = N - \lambda/\mu$, where μ^{-1} is the average trip duration.

Remark 1. *To ensure stability of the queue the model requires $N_I > 0$, i.e., $N > \lambda/\mu$. This is consistent with the TNC business model that "relies upon very short wait times for passengers requesting rides, which in turn depends on a large supply of available but idle drivers and vehicles" [2]. For instance, New York has an average of 5089 TNC vehicle [15], 187 passenger per minute, and a trip takes 16.3 minutes, i.e., $\mu = 1/16.3$ min^{-1}. This gives $N_I = N - \lambda/u = 2041$. Cities with limited supply of drivers (as in US suburbs and in cities like Singapore) require a distinct model [18], [27].*

3.2. Passenger and Driver Incentives

The passenger arrival rate λ and the number of drivers N are endogenously determined in the market equilibrium.

Passenger Incentives: Passengers choose their rides from available options like app-based TNCs, public transit, walking, or biking, by comparing their prices and waiting times. We model the cost of the app-based ride-hailing service as

$$c = \alpha t_w + \beta p_f, \tag{2}$$

where t_w is the average waiting time (from sending a request to being picked up), p_f is the per trip fare of the ride-hailing service [6], and α and β specify the passenger's trade-off between convenience and money. We refer to c as the total cost of a TNC trip, including the trip fare plus the money value of the trip time.

In the ride-hailing service, a ride is initiated when a passenger sends a request to the platform, and is completed when the passenger is dropped off at the destination. We divide a ride into three periods: (1) from the ride request being received to a vehicle being assigned; (2) from a vehicle being assigned to passenger pickup; (3) from passenger pickup to passenger drop-off. Let t_m, t_p and t_o be the average duration of these periods. Here t_m is the average waiting time in the M/G/N queue. Assuming that the platform matches the passenger to the nearest idle vehicle, t_p depends on the distance of the nearest idle vehicle to the passenger. Typically, t_m ranges from a couple of seconds to a half minute, and t_p is between three to six minutes. The sum of t_m and t_p is the passenger waiting time, denoted as $t_w = t_m + t_p$. Clearly, the passenger waiting time depends on the average number of idle vehicles N_I. We denote t_w as a function of N_I, i.e., $t_w(N_I)$, and impose the following assumption

[6] Most app-based ride-hailing platforms charge passengers based on the formula: total cost = base fare + price/mile × trip miles + price/time × trip time. p_f represents the sum of these three costs.

Assumption 1. *The function $t_w : \mathbb{R}_+ \to \mathbb{R}_+$ is convex, decreasing and twice differentiable.*

This assumption says that passenger waiting time decreases as N_I increases, and the marginal benefit of recruiting extra vehicles to reduce waiting time diminishes as N_I increases. It is a standing assumption throughout the paper.

In some special cases, the waiting time function $t_w(\cdot)$ can be derived analytically. Let $d(x)$ denote the distance of a passenger requesting a ride at location x in a city to the nearest idle vehicle. Let N_{I_0} be the average number of idle TNC vehicles before regulatory intervention (more precisely specified later). We have the following proposition.

Proposition 1. *Consider a city with an arbitrary geometry . Assume that (1) the platform matches each arriving passenger to the nearest idle vehicle available; (2) idle vehicles are uniformly and independently distributed across the city; (3) location of passengers requesting a ride is uniformly distributed across the city, independently of the position of idle vehicles. Then,*

$$\mathbb{E}_x\{d(x)|N_I \textit{ idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_x\{d(x)|N_{I_0} \textit{ idle vehicles}\}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}, N_{I_0}^{\frac{1}{8}}N_I^{-\frac{9}{8}}, N_I^{-1}\}\right)\right). \quad (3)$$

The result implies that the average pickup time t_p is (approximately) inversely proportional to the square root of the number of idle vehicles since $t_p = \frac{\mathbb{E}_x\{d(x)|N_I \,\text{idle vehicles}\}}{v}$, with v being the average traffic speed. The result recalls Mohring's "square root law" [28] and can be explained intuitively as follows. Consider a square city of unit size with N_I idle vehicles located in a grid with each idle car equally distant from its four closest neighbors to its left, right, top, or bottom, then the shortest distance between idle cars is equal to $\frac{1}{\sqrt{N_I}}$. The exact proof of Proposition 1 for a city of general shape and when the locations of idle vehicles and passengers are random is deferred to Appendix A.

Remark 2. *Proposition 1 has a few limitations. First, the platform may wait to accumulate idle vehicles and waiting passengers before matching [29]. This can potentially benefit the passenger/driver as they receive a closer match after waiting for a few more seconds. We do not capture this in Proposition 1. Second, we assume that both passengers and vehicles are uniformly and independently distributed across the city. In practice, passengers/drivers may strategically choose their locations to wait for the next vehicle/customer. This is also not considered in Proposition 1. Nevertheless, we emphasize that our analysis does not require any specific form of function $t_w(\cdot)$. The qualitative results of this paper hold as long as Assumption 1 is satisfied, and the result of Proposition 1 is only used to generate the numerical results.*

In (3) we select N_{I_0} as a reference so that $\frac{\mathbb{E}_x\{d(x)|N_{I_0} \,\text{idle vehicles}\}}{v}$ can be computed from available TNC data for a city. For instance, on average, Manhattan has 5089 TNC vehicles on the road. Every minute there are 187 new TNC trips. Each trip takes around 16.3 minutes [15], and passengers wait 5 minutes for pickup [30]. In this case we have $N = 5089$, $\lambda = 187$ trips/min, $\mu = 1/16.3$ min^{-1}, and the average pickup time is $\simeq 5$ min. Taking $N_{I_0} = N - \lambda/\mu \simeq 2041$, then $\frac{\mathbb{E}_x\{d(x)|N_{I_0} \,\text{idle vehicles}\}}{v} \simeq 5$min, and the pickup time function (3) becomes:

$$\mathbb{E}_x\{d(x)|N_I \,\text{idle vehicles}\} = \frac{226}{\sqrt{N_I}}\left(1+\mathcal{O}\left(\max\{2041^{-1}, 2.59N_I^{-\frac{9}{8}}, N_I^{-1}\}\right)\right),$$

The estimate in Proposition 1 has approximation error $\mathcal{O}(N_{I_0}^{-1} + (1+(\frac{N_{I_0}}{N_I})^{\frac{1}{8}})N_I^{-1})$ for large N_I and N_{I_0}. Medium to big sized cities usually have a few thousands of TNC vehicles[7], so (3) is a good approximation to the average pickup time for practical parameter values. In summary, we have:

[7]San Francisco has around 6000 active TNC vehicles on average [13], and Manhattan has more than $10,000$ TNC vehicles during peak hours [15].

Corollary 1. *Assume that all the conditions in Proposition 1 hold, and the ride confirmation time t_m is negligible compared to t_p, i.e., $t_w = t_m + t_p \simeq t_p$. Then*

$$t_w \simeq t_p = \frac{1}{v}\mathbb{E}_x\{d(x)|N_I \, idle\, vehicles\} \simeq \frac{M}{\sqrt{N_I}}, \tag{4}$$

where $M = \frac{1}{v}\sqrt{N_{I_0}}\mathbb{E}_x\{d(x)|N_{I_0}^{-1} \, idle\, vehicles\}$ and $N_I = N - \lambda/\mu$.

The ride-hailing platform has a distinctive supply-side *network externality*. As the number of drivers increases, so do the number and spatial density of idle drivers which, in turn, reduces pickup time and increases service quality. This enables larger platforms to offer the same service quality at a lower cost. For instance, consider a small platform and a large platform with the same vehicle occupancy. Assume the small platform is half the size of the large platform in terms of number of vehicles and passengers. Based on (4), the waiting time for the small platform, t_w^s, is $\sqrt{2}$ times that of the large platform, t_w^l. Let $t_w^s = 6$ min, then we have $t_w^l = 4.2$ min. The monetary value of this difference is $\alpha * (t_w^s - t_w^l) = \$5.7$ (See Section 4.2 for the value of α). This indicates that the smaller platform has to lower the rider fare by \$5.7 to attract the same number of passengers of the larger platform.

Remark 3. *The ride confirmation time t_m is orders of magnitude smaller than the pickup time t_p in large cities in the US. For instance, New York city has an average of 5089 TNC vehicles, 187 passengers per minute, and each trip takes 16.3 minutes [15]. If passenger arrivals are Poisson, then t_m as the average waiting time in the M/G/N queue is sub-second (virtually 0). In areas with limited supply of drivers, t_m could be significant and can not be neglected. We can add t_m to the travel cost. We conjecture that in this case if $t_m + t_w$ satisfies Assumption 1 the qualitative results of the paper still hold.*

Passengers have a reservation cost that summarizes their other travel options: if the TNC travel cost c is greater than the reservation cost, the passenger abandons the TNC for an alternative transport mode. We assume that the reservation costs of passengers are heterogeneous, and let $F_p(c)$ be the cumulative distribution of reservation costs. The passenger arrival rate then is given by

$$\lambda = \lambda_0 \left[1 - F_p\bigg(\alpha t_w(N - \lambda/\mu) + \beta p_f\bigg)\right] \text{ rides/min}, \tag{5}$$

where λ_0 is the arrival rates of potential passengers total travel demands in the city. Note that the trip time t_o does not depend on λ or N, so we absorb it into F_p as a constant. According to (5), the passengers that use the app-based ride-hailing service are all potential passengers except those that leave the platform because its cost is greater than their reservation cost.

Driver Incentives: Drivers are sensitive to earnings and respond to the offered wage by joining or leaving the platform. The average hourly wage of TNC drivers is

$$w = \frac{\lambda p_d}{N}, \tag{6}$$

wherein p_d is the per trip payment the driver receives from the platform. The platform keeps the difference between p_f and p_d as its commission or profit. In 2018 Uber collected \$41B from passengers of which drivers received 78% corresponding to a 22% commission rate [10], and Lyft collected \$8B from passengers and received 26.8% as commission [31].

The average hourly wage (6) is derived as follows. The total platform payment to all drivers sums to λp_d \$/min. Therefore the average hourly wage per driver is $\lambda p_d \times 60$ \$/hr divided by N, where the constant 60 captures the time period of one hour.

Each driver has a reservation wage. He joins TNC if the platform wage is greater than his reservation wage. We assume that the reservation wages of drivers are heterogeneous, and denote $F_d(c)$ is the cumulative distribution of reservation wages across the population of drivers. Hence

$$N = N_0 F_d\left(\frac{\lambda p_d}{N}\right). \tag{7}$$

Here N_0 is the number of potential drivers (all drivers seeking a job). For ease of notation we drop the constant factor 60 from the hourly wage formula and absorb it in the function F_d in (7). According to (7), the number of TNC drivers is the number of potential drivers multiplied by the proportion that joins the platform since their reservation wage is smaller than w.

Remark 4. *In practice, both supply and demand of a ride-hailing system vary within a day. This can be approximated in a quasi-static analysis by varying λ_0 and N_0 for peak and off-peak hours.*

4. TNC decisions in absence of regulation

The objective of the app-based ride-hailing platform adapts over time. In the initial phase it subsidizes passengers and drivers to grow the business. Eventually it shifts to maximizing the profit. Here we focus on profit maximization assuming that the platform is unregulated. Platform subsidy and competition are discussed in Section 8.

4.1. Pricing without Regulation

The platform rent or profit is

$$\Pi = \lambda(p_f - p_d). \tag{8}$$

In a certain period (e.g., each minute), λ trips are completed. Since the platform pockets $p_f - p_d$ from each trip, the total rent in this period is (8).

In the absence of regulation, the platform maximizes its rent subject to the market equilibrium conditions:

$$\max_{p_f \geq 0, p_d \geq 0} \lambda(p_f - p_d) \tag{9}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\Big(\alpha t_w(N - \lambda/\mu) + \beta p_f\Big)\right] & \text{(10a)} \\ N = N_0 F_d\left(\dfrac{\lambda p_d}{N}\right) & \text{(10b)} \end{cases}$$

We have the following result on the existence of solution to (10):

Proposition 2. *If $F_p(\alpha t_w(N_0)) < 1$ and $N_0 > \lambda_0/\mu$, there exist strictly positive λ, N, p_f and p_d that constitute a market equilibrium satisfying (10).*

The proof can be found in Appendix B. The assumption $F_p(\alpha t_w(N_0)) < 1$ means that when the platform recruits all potential drivers N_0 and sets the ride price at 0 ($p_f = 0$), there will be a positive number of passengers. This assumption rules out the situation in which passenger reservation costs are so low and driver reservation wages are so high that supply and demand curves do not intersect.

Since (9) is not a convex problem, it is not straightforward to determine its solution. One approach is via numerical computation as in [20]. This is suitable for small problems. Instead, we proceed analytically.

We view $\lambda = \lambda(p_f, p_d)$ as a function of p_f and p_d determined implicitly by (10). The first order necessary conditions of (9) then simplify to

$$\begin{cases} \dfrac{\partial \lambda}{\partial p_f}(p_f - p_d) + \lambda = 0 & \text{(11a)} \\ \dfrac{\partial \lambda}{\partial p_d}(p_f - p_d) - \lambda = 0 & \text{(11b)} \end{cases}$$

in which (11a) is equivalent to $\dfrac{\partial \Pi}{\partial p_f} = 0$ and (11b) is $\dfrac{\partial \Pi}{\partial p_d} = 0$. For non-convex problems, these conditions are only necessary. However, they are sufficient in the following case.

Proposition 3. *Assume that (a) the waiting time function t_w satisfies (4); (b) the reservation cost and the reservation wage are uniformly distributed as $F_p(c) = \min\{e_p c, 1\}$ and $F_d(w) = \min\{e_d w, 1\}$, with $e_p \in \mathbb{R}$ and $e_d \in \mathbb{R}$; (c) the profit maximizing problem (9) has at least one solution at which the objective value is positive. Then the following equations have a unique solution (p_f, p_d, λ, N)[8], which is the globally optimal solution to (9).*

$$\begin{cases} \dfrac{\partial \lambda}{\partial p_f}(p_f - p_d) + \lambda = 0 & \text{(12a)} \\ \dfrac{\partial \lambda}{\partial p_d}(p_f - p_d) - \lambda = 0 & \text{(12b)} \\ \lambda = \lambda_0 \left[1 - F_p\left(\dfrac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f \right) \right] & \text{(12c)} \\ N = N_0 F_d\left(\dfrac{\lambda p_d}{N} \right) & \text{(12d)} \end{cases}$$

The proof of Proposition 3 is deferred to Appendix C. It asserts that (9) can be effectively computed by finding the unique solution to (12). Note that if the assumptions of Proposition 3 are not satisfied, we can still solve (9) by brute-force computation.

4.2. Numerical Example

We present a numerical example and calculate the platform's profit-maximizing decision (9). To apply Proposition 3, we assume that the waiting time function t_w satisfies (4), and that the reservation cost of passengers and the reservation wage of drivers are both uniformly distributed. Below we specify the model parameters used in the simulation.

Parameters: We take the TNC data for the Manhattan Central Business District (CBD) in New York city. It records all trips that started from or ended in Manhattan CBD on a regular weekday. Let L denote the average TNC trip distance. We obtain the following estimates based on [15]:

$$N = 5089,\ \lambda = 187 \text{ ride/min},\ L = 2.4 \text{ mile},\ t_o = 16.3 \text{ min},\ p_f = \$17/\text{trip},\ p_d = \$10.2/\text{trip}. \tag{13}$$

Note that t_o denotes the average TNC trip duration.

Our estimation proceeds as follows. Based on [15], each day TNC vehicles make 202,262 trips over 91,608 vehicle hours and 802,135 miles. On average, the vehicle are occupied 60% of the time [15]. Since there are

[8]Since $p_f = p_d = \lambda = N = 0$ is always a solution, throughout the paper by 'solution' we refer to strictly positive solutions unless otherwise stated.

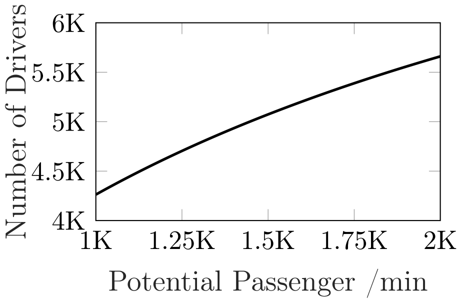

Figure 2: Number of drivers under different potential passengers.

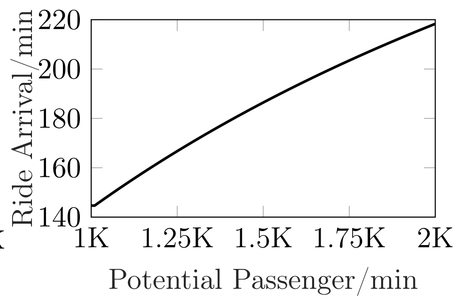

Figure 3: Arrival rates of Passengers (per minute).

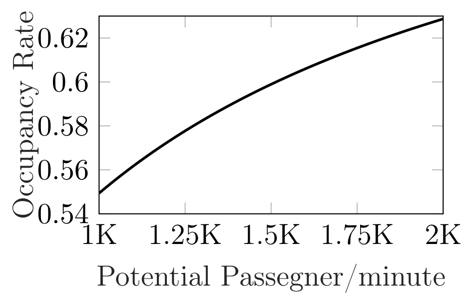

Figure 4: Occupancy rate under different potential passengers.

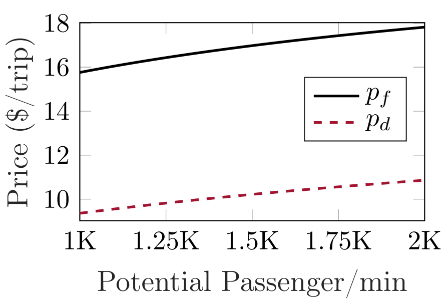

Figure 5: Per mile ride price and driver payment.

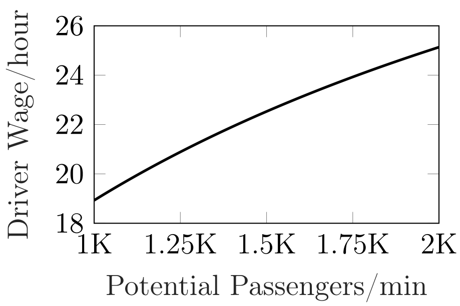

Figure 6: Driver wage per hour under different potential passengers.

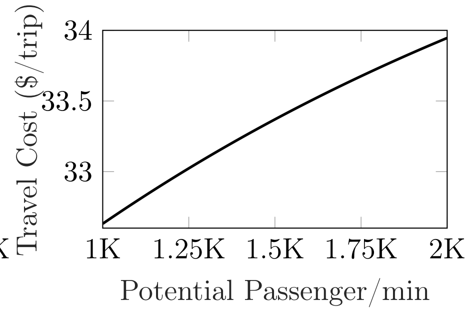

Figure 7: Passenger travel cost under different potential passengers.

virtually no rides between 1AM − 7AM, we divide the daily numbers by 18 (hours) to get $N = 91,608/18 = 5089$ and $\lambda = 202,262/18/60 = 187$ ride/min. The average trip length is

$$\frac{\text{total mileage}}{\text{number of trips}} \times \text{occupancy} = \frac{802,135}{202,262} * 0.6 = 2.4 \text{ mile}.$$

The average trip duration is

$$\frac{\text{vehicle hours}}{\text{number of trips}} \times \text{occupancy} = \frac{91,608}{202,262} * 0.6 * 60 = 16.3 \text{ min}.$$

We estimate that a 16-min, 2.4-mile ride In Manhattan costs \$17. TNC drivers in New York earn an average of \$22.6 per hour before expenses [2]. This suggests $w = \lambda p_d/N = \$22.6$, and so $p_d = 22.6N/\lambda = \$10.2$ per trip.

Note that our estimates (13) are solutions to the profit-maximizing problem (9). We now utilize these solutions to 'reverse-engineer' the model parameters $(N_0, \lambda_0, \alpha, \beta)$. In particular, we select $(N_0, \lambda_0, \alpha, \beta)$ so that the solutions to (9) match the real data (13). This can be realized by substituting (13) into (12) and solving the first-order conditions (12). We obtain:

$$N_0 = 13,512,\ \lambda_0 = 1512/\text{min},\ \alpha = 3.2\$/\text{min},\ \beta = 1,\ e_p = 0.0262,\ e_d = 1, M = 226, \mu = 1/16.3\text{min}^{-1}. \tag{14}$$

Empirical study suggests that value of travel time (VOT) in the range \$20 to \$100 per hour [32] and value of waiting time at 2 to 3 times that of in-vehicle travel time [33]. Our estimate of $\alpha = \$3.2$ per min corresponds to a VOT between \$64 and \$96 per hour.

Results: We vary λ_0 between 1000 and 2000 to study how the platform decision varies at different times of the day (λ_0 is large during peak hours). The results are shown in Figures 2-7. As λ_0 increases, the number of passengers (λ) and drivers (N) both increase. At the same time, occupancy rate (Figure 4), and the ride price increase (Figure 5). At peak hours, the drivers benefit since they earn more (Figure 6), but the passengers travel at a higher cost (2) due to the increased trip fare.

Note that as the number (λ_0 of potential passengers doubles from 1000 to 2,000 riders per minute, the profit-maximizing fare (p_f) per ride set by the platform increases by 13 percent from \$15.8 to \$17.8 per trip, driver payment (p_d) increases by 16 percent from \$9.4 to \$10.9 per trip, and the platform's share increases by 8 percent from \$6.4 to \$6.9. The 33 percent increase in driver wages from \$17.8 to \$25.1 per hour is due jointly to the increases in per trip payment to the driver and the vehicle occupancy (from 0.55 to 0.63). By the same token, a driver's hourly wage declines by 33 percent from peak to off-peak hours. Thus in the absence of a wage floor, drivers bear most of the risk of shifts in demand.

5. TNC decisions with wage floor

This section is devoted to platform pricing with a wage floor. A driver minimum wage w imposes the constraint $\lambda p_d/N \geq w$[9]. After a wage floor is imposed the platform may find it prohibitive to hire all drivers who wish to join and, thus, limit the entry of new drivers. We capture this by relaxing (10b) to the inequality (16b). The profit maximizing problem subject to a wage floor is

$$\max_{p_f \geq 0, p_d \geq 0, N} \lambda(p_f - p_d) \tag{15}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\left(\alpha t_w(N - \lambda/\mu) + \beta p_f \right) \right] & \text{(16a)} \\ N \leq N_0 F_d \left(\dfrac{\lambda p_d}{N} \right) & \text{(16b)} \\ \dfrac{\lambda p_d}{N} \geq w & \text{(16c)} \end{cases}$$

Constraint (16b) indicates that the platform can hire up to the number of all available drivers. This introduces an additional decision variable N. It can be solved via numerical computation as in [20]. Similarly to Proposition 2, we can show that (16) has at least one non-trivial solution if $F_p(\alpha t_w(N_0)) < 1$.

5.1. A Cheap-Lunch Theorem

Example: Consider an example for which we calculate the profit-maximizing prices (15) for different wage floors w. We assume that the waiting time function t_w is of form (4), and that the reservation cost of passengers and the reservation wage of drivers are both uniformly distributed. We set the model parameters as (13) and (14). We emphasize that these assumptions are only needed for numerical simulations. Our analysis does not depend on these assumptions or model parameters. Figures 8-16 reveal the market response to different levels of the wage floor, including number of drivers, arrival rates of passengers, vehicle occupancy rate, driver wage, passenger pickup time, platform prices, and platform profit. The response has three distinct regimes:

- $w <$ \$22.6: the wage floor constraint (16c) is inactive and the solution to (15) is the same as that to (9) because even in the absence of the minimum wage constraint the platform sets $w =$ \$22.6 to attract enough drivers.

[9] New York City Taxi and Limousine Commission imposes the minimum driver payment (1). We assume a constant speed, so per-minute price can be transformed to per-mile price, and we use p_d to represent the sum of the first and second term in (1). If we neglect the constant shared ride bonus, (1) is proportional to $\lambda p_d/N$.

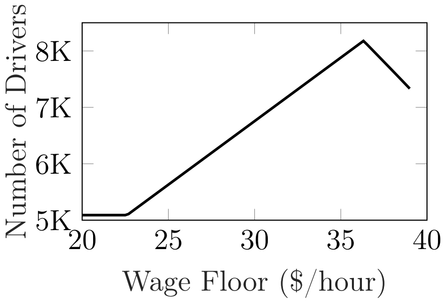

Figure 8: Number of drivers under different wage floors.

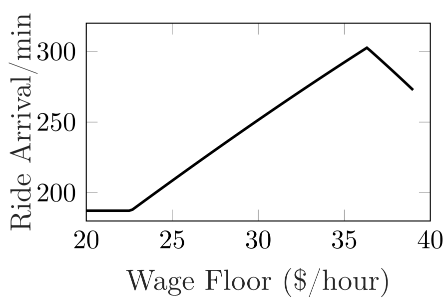

Figure 9: Arrival rates of Passengers (per minute).

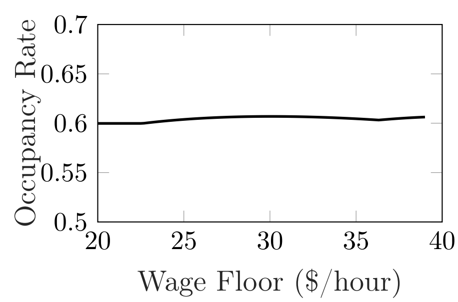

Figure 10: Occupancy rate under different wage floors.

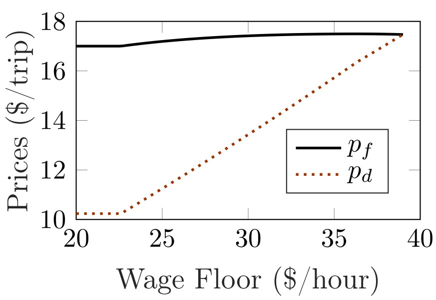

Figure 11: Per mile ride price and driver payment.

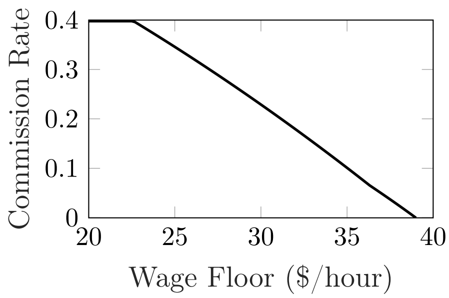

Figure 12: Commission rate defined as percentage of service fee in p_f.

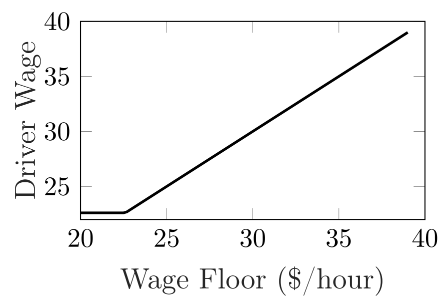

Figure 13: Driver wage per hour under different wage floors.

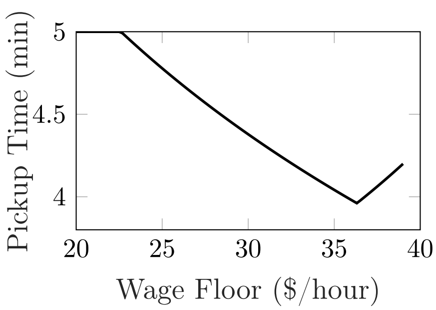

Figure 14: Passenger pickup time under different wage floor.

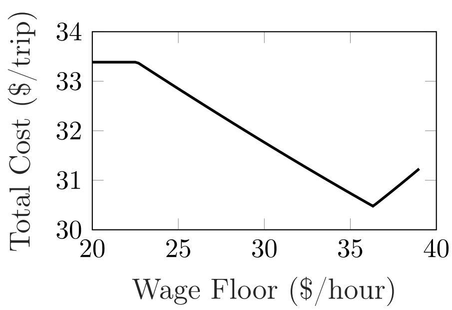

Figure 15: Total cost of passengers under different caps.

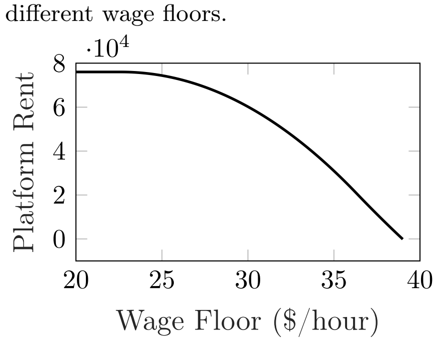

Figure 16: Platform profit ($/hour) under different wage floors.

- $22.6 $\leq w <$ $36.3: both (16b) and (16c) constraints are active. As the minimum wage increases, the platform hires all drivers whose reservation wage is below the minimum wage, the ride cost (2) goes down, the quality of service (pickup time) improves, driver wage increases, more passengers are served, and the platform profit reduces.

- $w \geq$ $36.3: only the wage floor constraint (16c) is active. As the minimum wage exceeds $36.3, the platform hires fewer drivers than wish to work, both ride fare (p_f) and pickup time (t_p) increase, fewer passengers take the ride-hailing option, the drivers who are hired earn more, and the platform profit reduces further.

According to Theorem 1 below, the qualitative behavior of most variables, including number of drivers, arrival of passengers, driver wage, travel cost, and platform rent remains consistent with Figures 8-16 for all model parameters. The behavior of p_f may depend on model parameters. When α is small, the trip fare p_f may decrease in the second regime of Figure 11. This is because for small α passengers are more sensitive to trip fare, and the platform may find it more effective to attract passengers by reducing the trip fare. However, we emphasize that the total travel cost (2) as the sum of p_f and pickup time always decreases in the second regime.

Remark 5. *When the wage floor reaches \$39 per hour, the platform profit is 0. In this case, the platform may exit the market. This regulatory risk associated with the ride-sharing business model is explicitly called out by Uber and Lyft in their IPO registration statements [10, 31]. In practice, it is unlikely that regulations will drive platform revenue to zero. New York city's wage floor of \$27.86 per hour (before expenses) will predictably lead to 10.5% decrease in platform profit from \$76K to \$68K per hour.*

Assuming the optimum solution to (15) is unique, write it as a function of w: $(N^*(w), \lambda^*(w), p_f^*(w), p_d^*(w))$. We have the following theorem.

Theorem 1. *Assume that (15) has a unique solution.*[10] *For any parameters $(N_0, \lambda_0, \alpha, \beta)$ and any distributions $F_p(\cdot)$ and $F_d(\cdot)$ that satisfy $F_p(\alpha t_w(N_0)) < 1$ and $N_0 > \lambda_0/\mu$, we have $\nabla_+ N^*(\tilde{w}) > 0$ and $\nabla_+ \lambda^*(\tilde{w}) > 0$, where ∇_+ denotes the right-hand derivative, and $\tilde{w}$ is the optimal driver wage in absence of regulation, i.e., the solution to (9).*

The proof of Theorem 1 can be found in Appendix D. Theorem 1 holds for every pickup time t_w that satisfies Assumption 1 and it does not assume any specific formula for pickup time or a specific matching algorithm utilized by the platform. This implies that the second regime always exists: when $w \le \tilde{w}$, the minimum floor constraint (16c) is inactive, so the solution is in regime 1. When $w = \tilde{w}$, the right-hand derivative of N and λ are both strictly positive. This corresponds to the beginning of the second regime, where the platform hires more drivers and serves more passengers. The increase in the number of drivers and passengers implies that the wage of drivers increases and the total cost of passengers decreases.

Discussion: The effect of minimum wage on labor markets has been the subject of many studies in labor economics since its inception as part of Fair Labor Standard Act of 1938, and it still remains a contentious topic among economists. We provide a brief overview of the existing literature on the effect of the minimum wage regulation on employment. We then discuss how our result connects to the current literature.

There is mixed empirical evidence on whether imposing a minimum wage has a positive or negative effect on employment; for instance, the authors in [34, 35] and [36] use the data from fast-food industry in Pennsylvania and New Jersey and draw drastically different conclusions on the impact of an increase in the minimum wage. A recent meta-study [37] found that it is equally likely to find positive or negative employment effect of the minimum wage in the literature; a similar observation is made in [38].[11]

A similar division in economic theory literature exists regarding the direction of the employment effect of minimum wage [41]. One strand of work that assumes that the labor market is perfectly competitive concludes that minimum wage has a negative effect on employment. Another strand of work considers a monopsony framework where the employer has bargaining power over the wage, and the labor demand is upward-sloping. This work contends that the minimum wage may actually increase employment [42].

Our results above are similar to those of the monopsony framework in the literature. In a ride-sharing market, a TNC has market power over drivers as it explicitly sets price p_d for drivers. Moreover, given the significant size of its drivers (e.g. Uber is the largest for-profit employer in New York city if we consider drivers as employees [2]), TNC does not face a perfectly competitive labor market. [12]

Our model is different from the standard economic equilibrium model in which supply and demand are equal. We consider a model where the supply (drivers N) must exceed the demand (riders λ), and the

[10] For almost all parameter values there will not be multiple solutions with the same optimal value.

[11] We refer the interested reader to [39, 40] for surveys of studies on the effects of minimum wage.

[12] We note that in labor economics, by monopsony they do not only refer to the traditional *company town* with a single employer with full market power. The term monopsony applies more broadly to cases where the employer has some market power to set wages and faces upward-sloping labor supply [43].

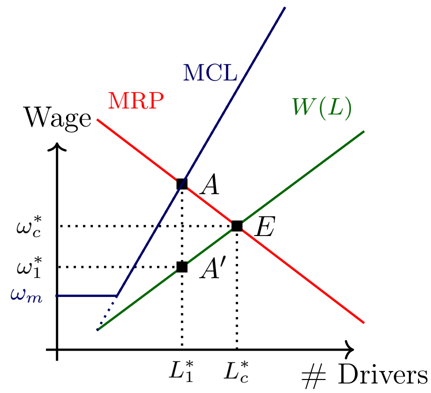

Figure 17: Regulated market with wage floor $w_m < \omega_1^*$ (first regime)

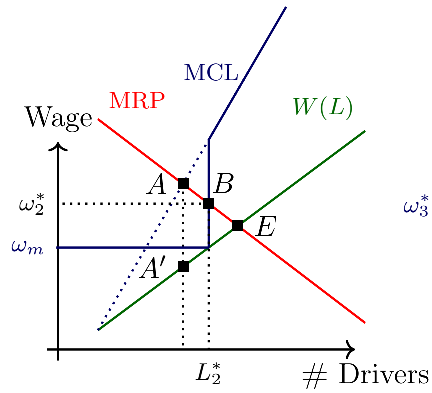

Figure 18: Regulated market with wage floor $\omega_1^* \leq w_m \leq \omega_c^*$ (second regime)

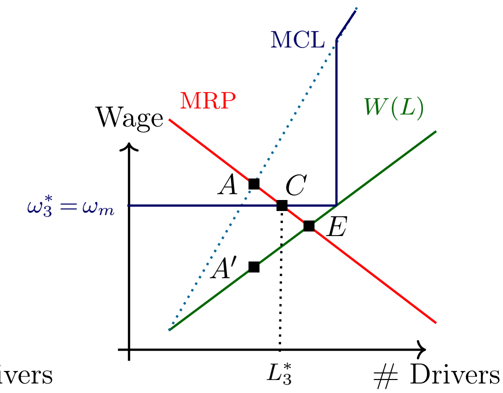

Figure 19: Regulated market with wage floor $w_m > \omega_c^*$ (third regime)

difference between the supply and demand (idle cars N_I) contributes to the total cost the riders faces through waiting times $t_w(N_I)$. Nevertheless, we can use the monopsony framework to provide an intuitive explanation of Theorem 1 below.

Consider Figure 17, where curve $W(L)$ depicts the wage corresponds to every employment level L and curve MRP represents the resulting marginal-revenue product equivalent to employment level L. We note that in deriving the MRP curve, we ignore the effect of waiting time t_w and assume that $c = \beta p_f$. The intersection of $W(L)$ and MRP (point E) determines the outcome in a competitive labor market with employment L_c^* and wage ω_c^*. However, a TNC does not face a perfectly competitive labor market, and sets wages to maximize its profit. From $W(L)$ we can determine the marginal cost of labor MCL defined as the marginal cost the TNC has to pay to hire one more driver; we note that to hire an additional driver the TNC has to increase the wage for all of his existing drivers, thus, MLC curve lies above $W(L)$. Figures 17-19 depicts the resulting MCL curves for tree different regimes depending on the value of the minimum wage ω_m. The optimal employment level and wage can be determined by the intersection of MRP and MCL curves, i.e. point A in the first regime, point B in the second regime, and point C in the third regime.

As the minimum wage ω_m increases, it is easy to verify that the number of drivers is constant in the first regime (Figure 17), increases in the second regime (Figure 18), and decreases in the third regime (Figure 19). Ignoring the effect of idle vehicles and waiting time t_w on cost c, the number of riders follow a similar pattern as the number of drivers. Consequently, the cost for riders is constant in the first regime, decreases in the second regime, and increases in the third regime. The above monopsony argument does not capture the presence of idle vehicles and their effect on waiting time t_w. The results of Theorem 1 establishes the result formally incorporating the impact of t_w on total cost c to riders.

6. TNC decisions with cap on number of drivers

Let N_{cap} be the cap on the total number of drivers. With a cap constraint, the platform pricing problem is

$$\max_{p_f \geq 0, p_d \geq 0} \lambda(p_f - p_d) \tag{17}$$

$$
\begin{cases}
\lambda = \lambda_0 \left[1 - F_p\Big(\alpha t_w(N - \lambda/\mu) + \beta p_f\Big)\right] & \text{(18a)} \\
N = N_0 F_d \left(\dfrac{\lambda p_d}{N}\right) & \text{(18b)} \\
N \leq N_{cap} & \text{(18c)}
\end{cases}
$$

It is unnecessary to relax (18b), since the platform can always lower p_d to increase its profit. As with Proposition 2 we can show that (18) admits a non-trivial solution if $F_p(\alpha t_w(N_0)) < 1$. This is a non-convex program. One approach is via numerical computation as in [20]. This is suitable for a small problem. Alternatively, we can find the optimal solution based on first order conditions for the following special case:

Proposition 4. *Assume that (a) the waiting time function t_w satisfies (4); (b) the reservation cost and the reservation wage are uniformly distributed as $F_p(c) = \min\{e_p c, 1\}$ and $F_d(w) = \min\{e_d w, 1\}$, with $e_p \in \mathbb{R}$ and $e_d \in \mathbb{R}$; (c) the profit maximizing problem (17) has at least one solution at which the objective value is positive. Then the first order conditions of (17) admit a unique solution (p_f, p_d, λ, N), which is the globally optimal solution to (17).*

The proof is deferred to Appendix E, and the first order conditions of (17) are defined in the proof. **Example:** Consider an example where the platform solves the profit-maximizing problem (17) for different levels of N_{cap}. We assume that the waiting time function t_w is of form (4), and that the reservation cost of passengers and the reservation wage of drivers are both uniformly distributed. We set the model parameters as (13) and (14).

Figures 20-28 exhibit the market response to different caps on the total number of vehicles. These responses include the arrival rates of passengers, occupancy rate, platform prices, driver wage, pickup time and platform profit. It is more instructive to "read" the figures from right to left, as the cap decreases. As the cap decreases, the supply of vehicles drops (Figure 20), so it is more difficult for passengers to find a ride. In this case, pickup time increases (Figure 26), and the number of rides decreases (Figure 21). Here are some interesting observations:

- The platform loses passengers faster than it loses drivers. This is evidenced by the drop in occupancy (Figure 22).
- The pickup time increases at an increasing rate (Figure 26). This is just a counterpart of the aforementioned network externality.
- Both trip fare and driver wage drop (Figures 23, 25).

These observations can be explained. As the cap reduces the number of drivers, the passenger pickup time increases. Since t_p is a decreasing convex function of N, it has an increasing derivative as N decreases. Therefore customers leave the platform at an increasing rate as N_{cap} decreases. This rate is greater than the decreasing rate of N_{cap}, so occupancy rate decreases. In this case, the platform loses passengers quickly, and has to reduce trip prices to keep passengers from leaving. This further squeezes driver pay (Figure 25).

A surprising fact is that the cap on number of drivers hurts the earning of drivers (Figure 25). This is contrary to the common belief that limiting their number will protect drivers, as expressed in the regulation freezing the number of TNC vehicles in New York for one year [16]. This happens because the platform hires drivers with lowest reservation wage first. That is, with a smaller cap on the number of drivers, the platform responds by reducing driver pay and hiring drivers with lower reservation costs. Thus the benefit of limiting supply is intercepted by the platform. This is in contrast with the situation of taxis that need a medallion to operate. A limit on the number of medallions will increase their value and benefit their

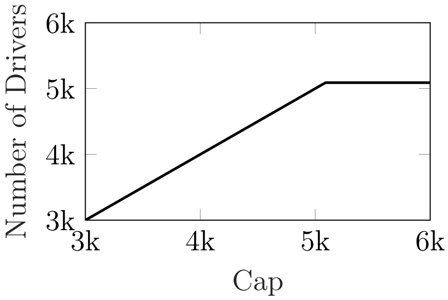

Figure 20: Number of drivers under different caps.

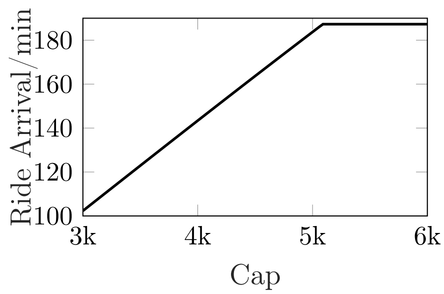

Figure 21: Arrival rates of Passengers (per minute).

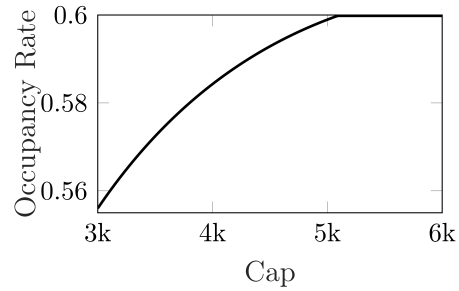

Figure 22: Occupancy rate under different caps.

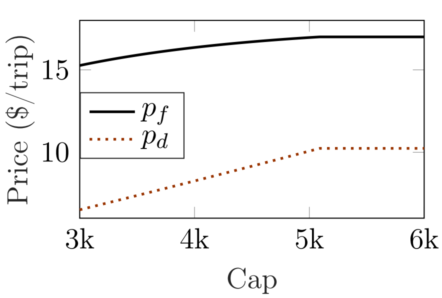

Figure 23: Per mile ride price and driver payment under different caps.

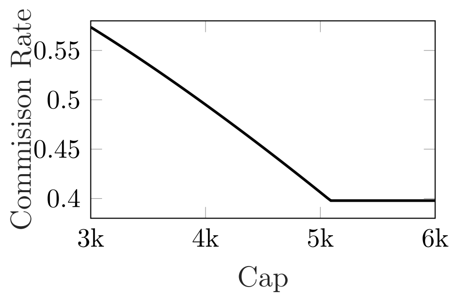

Figure 24: Commission rate defined as percentage of service fee in p_f.

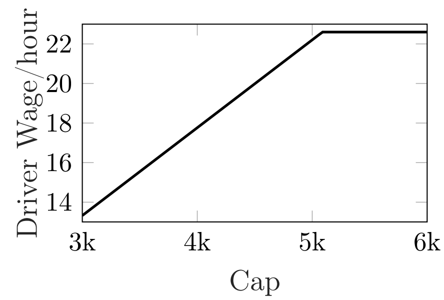

Figure 25: Driver wage per hour under different caps.

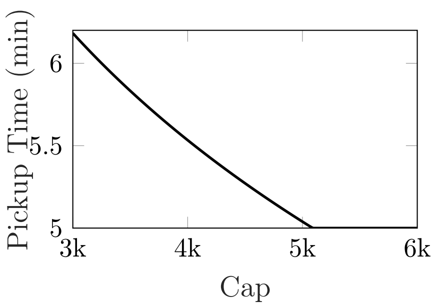

Figure 26: Passenger pickup time under different caps.

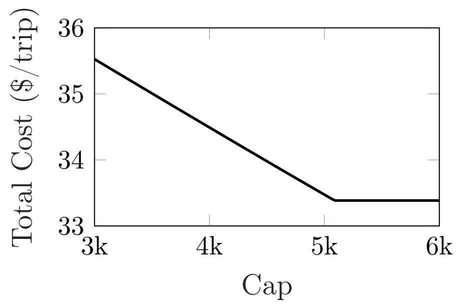

Figure 27: total cost of passengers under different caps.

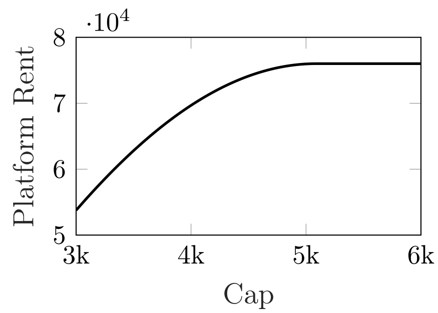

Figure 28: Platform profit ($/hour) under different caps.

owners[13], who may be taxi drivers. In the TNC case, the platform accumulates the increased value. This conclusion holds in general and is not affected by the model parameters.

7. TNC decisions with congestion surcharge

As of Jan 2019, all trips by *for-hire* vehicles that cross 96th street in NYC incur a congestion surcharge of \$2.75 per TNC trip. We model the likely impact of this policy by adding a congestion surcharge p_c to the travel cost (2). The profit-maximizing problem for the platform now is

$$\max_{p_f \geq 0, p_d \geq 0} \lambda(p_f - p_d) \tag{19}$$

[13]The platform revenue (sum of platform rent and driver payments) divided by the number of drivers increases as the cap decreases.

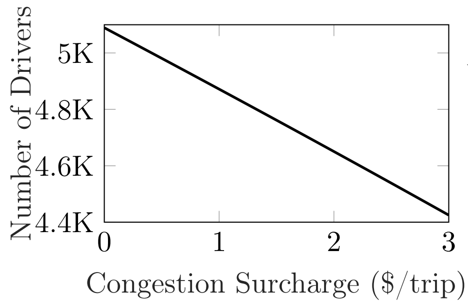

Figure 29: Number of drivers under different surcharge.

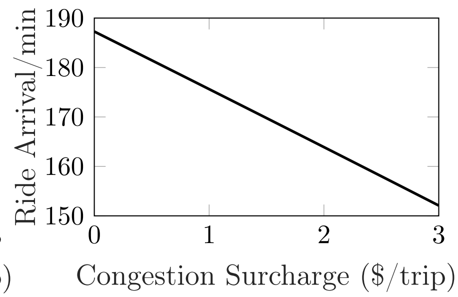

Figure 30: Arrival rates of Passengers (per minute).

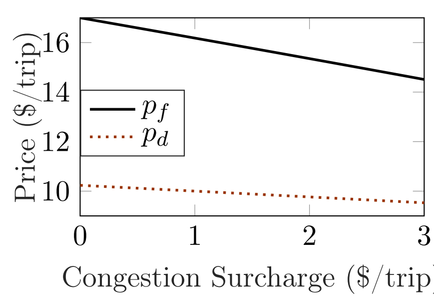

Figure 31: Per mile ride price and driver payment under different surcharge.

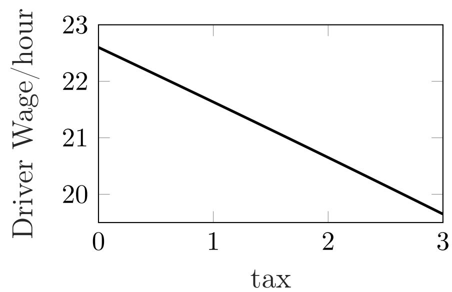

Figure 32: Driver wage per hour under different surcharge.

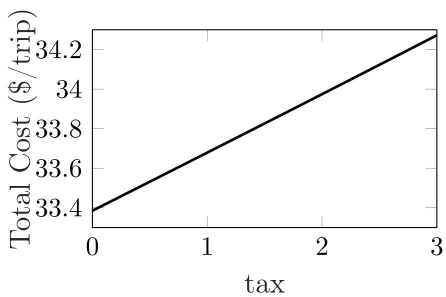

Figure 33: Total cost of passengers under different surcharge.

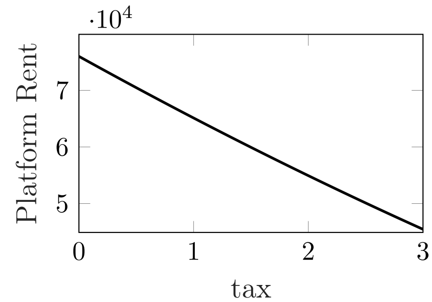

Figure 34: Platform profit ($/hour) under different surcharge.

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\left(\alpha t_w(N - \lambda/\mu) + \beta p_f + \beta p_c\right)\right] & \text{(20a)} \\ N = N_0 F_d\left(\dfrac{\lambda p_d}{N}\right) & \text{(20b)} \end{cases}$$

As with Proposition 2, we can show that the constraint set (20) is non-empty if $F_p\Big(\alpha t_w(N_0) + \beta p_c\Big) < 1$. Since (19) is not a convex program, one approach to solve (19) is via brute-force computation [20]. Alternatively, we can show that the first order conditions are sufficient for global optimization for some special cases:

Proposition 5. *Assume that (a) the waiting time function t_w satisfies (4); (b) the reservation cost and the reservation wage are uniformly distributed as $F_p(c) = \min\{e_p c, 1\}$ and $F_d(w) = \min\{e_d w, 1\}$, with $e_p \in \mathbb{R}$ and $e_d \in \mathbb{R}$; (c) the profit maximizing problem (19) has at least one solution at which the objective value is positive. Then the following equations have a unique solution (p_f, p_d, λ, N), which is the globally optimal solution to (19).*

$$\begin{cases} \dfrac{\partial \lambda}{\partial p_f}(p_f - p_d) + \lambda = 0 & \text{(21a)} \\ \dfrac{\partial \lambda}{\partial p_d}(p_f - p_d) - \lambda = 0 & \text{(21b)} \\ \lambda = \lambda_0 \left[1 - F_p\left(\dfrac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f + \beta p_c\right)\right] & \text{(21c)} \\ N = N_0 F_d\left(\dfrac{\lambda p_d}{N}\right) & \text{(21d)} \end{cases}$$

The proof is similar to that for Proposition 3, and is therefore omitted.

We estimate the platform's response to various values of congestion surcharge p_c by numerical simulation. In this example, we impose the same assumptions and model parameters as in Section 4.2. Simulation results, presented in Figure 29-34, show that under a congestion surcharge of \$2.75 per trip, the number of TNC vehicles drops by 11.9% from 5089 to 4480, TNC rides reduce by 17.1% from 187/min to 155/min, and platform revenue shrinks by 37.3% from \$76,006/hour to \$47,686 per hour. This also suggests that the TNC business model is vulnerable to regulatory risk.

8. Extensions

We formulate some extensions of the basic model to examine platform subsidy, platform competition and autonomous mobility on demand.

8.1. Platform Subsidy

The ride-hailing platform company is not always a short-term profit maximizer. In its early stages, it tries to grow its business via subsidies to both passengers and drivers. To model this, we consider a ride-hailing platform that sets prices to maximize the number of rides or passengers λ subject to a reservation revenue R, which may be positive or negative (negative R indicates subsidy):

$$\max_{p_f \geq 0, p_d \geq 0} \lambda \tag{22}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\left(\alpha t_w(N - \lambda/\mu) + \beta p_f\right)\right] & \text{(23a)} \\ N = N_0 F_d\left(\dfrac{\lambda p_d}{N}\right) & \text{(23b)} \\ \lambda(p_f - p_d) \geq R & \text{(23c)} \end{cases}$$

For notational convenience, let $(\lambda^\star, p_f^\star, p_d^\star)$ be the solution to (22), and denote $(\tilde{\lambda}, \tilde{p}_f, \tilde{p}_d)$ as the solution to the non-subsidy case (9).

We define subsidy as $\epsilon_f = p_f^\star - \tilde{p}_f$ and $\epsilon_d = \tilde{p}_d - p_d^\star$, where ϵ_f and ϵ_d represent the subsidy to passengers and drivers, respectively. Note that this definition essentially compares $(p_f^\star, p_d^\star)$ to the profit-maximizing prices. For ease of understanding, we define $B = \tilde{\lambda}(\tilde{p}_f - \tilde{p}_d) - R$ as the subsidy budget. When reservation revenue is the maximal profit, i.e., $R = \tilde{\lambda}(\tilde{p}_f - \tilde{p}_d)$, the subsidy budget is 0, and $\epsilon_f = \epsilon_d = 0$.

We estimate the platform's ridership under different levels of subsidy. In the numerical example we impose the same assumptions and model parameters as in Section 4.2. Simulation results presented in Figure 35-37 show that the platform should always subsidize both sides of the market, regardless of the subsidy level. Another interesting observation is that the platform should subsidize drivers more than it does passengers. This conclusion, however, depends on the elasticities of demand and supply because the platform has to grow both sides of the market to maximize profit. Under a fixed budget, the platform allocates more subsidy to the less price-sensitive side as it costs more to grow this side of the market by one unit.

8.2. Platform Competition

Consider two platforms (e.g., Uber and Lyft) competing with each other to maximize their profits. The profits are coupled through the market response to the joint decisions of both platforms: passengers choose the platform with lower overall cost, and drivers work for the platform with a higher wage rate. This subsection modifies the model to capture this competition.

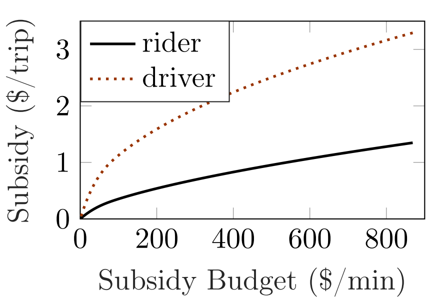

Figure 35: Subsidies to passengers and drivers under different subsidy budgets.

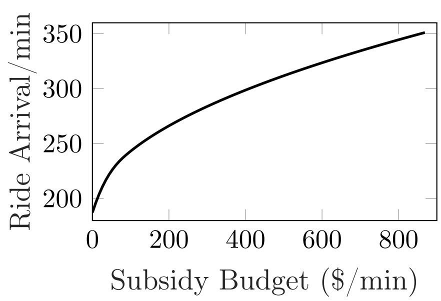

Figure 36: Arrival rate of passengers under different subsidy budgets.

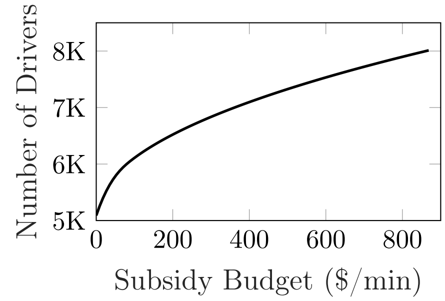

Figure 37: Number of drivers under different subsidy budgets.

Each platform selects its passenger fare and driver wage. Passengers and drivers respond to the platform prices until the market settles down. Assume that when the market settles down, both platforms survive with a positive profit. In this case neither passengers nor drivers deviate from their choice of platform at the market equilibrium, so the passenger costs and driver wages for the two platforms are equal. This gives rise to the following profit maximization problem for one platform, given the pricing decisions (p'_f, p'_d) of its competitor:

$$\max_{p_f \geq 0, p_d \geq 0} \lambda(p_f - p_d) \tag{24}$$

$$\begin{cases} \alpha t_w(N - \lambda/\mu) + \beta p_f = \alpha t_w(N' - \lambda'/\mu) + \beta p'_f & \text{(25a)} \\ \dfrac{\lambda p_d}{N} = \dfrac{\lambda' p'_d}{N'} & \text{(25b)} \\ \lambda + \lambda' = \lambda_0 \left[1 - F_p\Big(\alpha t_w(N - \lambda/\mu) + \beta p_f\Big)\right] & \text{(25c)} \\ N + N' = N_0 F_d\left(\dfrac{\lambda p_d}{N}\right) & \text{(25d)} \end{cases}$$

Constraints (25a) and (25b) guarantee that if both platforms have positive number of passengers and drivers, then the passenger cost and driver wage in the two platforms are the same, so no passenger or driver has an incentive to switch platforms. Note that the market outcomes $(N, \lambda, N', \lambda')$ are not given. Instead, they are governed by the market equilibrium conditions (25a)-(25d) and the platform prices (p_f, p_d, p'_f, p'_d). One difference between (24)-(25d) and the monopoly case (9)-(10b) is that in the former the waiting time for each TNC depends on the number of its own idle vehicles rather than on the sum of the idle vehicles of the two platforms. The second difference, by contrast, is that the wage rate (25d) is determined by the sum of the demand for drivers by both platforms.

Analogously, the second platform's decisions (p'_f, p'_d) will maximize its own profits given the decisions (p_f, p_d) of the first. The solution of the two decision problems will be a Nash equilibrium.

Due to non-convexity of (24), the question of existence and uniqueness of Nash equilibrium remains open. It is possible that the two platforms will split the heterogeneous passengers, with one platform offering a higher fare, lower waiting time, luxury rides to passengers with higher reservation cost; the emergence of such equilibrium outcomes with product quality differentiation was first demonstrated by [44, 45] in oligopolies.

8.3. Autonomous Vehicles

Autonomous vehicles (AV) will revolutionize road transportation [46]. AV companies claim they will banish 94 percent of all accidents attributed to human error [47]. So commuters can sit back and relax, work, or entertain themselves. Eventually there will be hardly any need for human drivers. The impact on our lives will be profound. Uber and Lyft have R&D efforts to build self-driving ride-hailing vehicles. Billions of venture capital are flowing into the race to develop AVs.

We model a company that owns and operates a fleet of autonomous vehicles to provide autonomous mobility on demand (AMoD) service [48, 49], and compare it to a ride-hailing service with human drivers. We modify the model (9)-(10b) to relate the decisions of an AMoD monopoly and those of a TNC monopoly. The AMoD monopoly will set its ride rates to maximize its profit (26) subject to demand (27):

$$\max_{p_f \geq 0, N \geq 0} \lambda(p_f - N c_{av}) \tag{26}$$

$$\text{s.t. } \lambda = \lambda_0 \left[1 - F_p\Big(\alpha t_w (N - \lambda/\mu) + \beta p_f\Big)\right] \tag{27}$$

Here N is the number of deployed AVs and c_{av} is the per ride investment and operating cost of an AV. Comparing this with the TNC decision making model (9)-(10b) we get a straightforward formal identification:

number of AVs deployed $\doteq$ number of drivers hired, and wage rate $w \doteq c_{av}$ vehicle cost.

Following the NYTLC ruling, we take $w = \$27.86$ per hour or an annual cost of \$55,000 for 2,000 hours per year of driver (plus vehicle) service. So for the AMoD monopoly to be as profitable as the TNC monopoly (26) implies that an AV's annual investment and operating cost should be smaller than \$55,000. How likely is this?

Today's AVs do not meet this cost target. In records submitted to the California Department of Motor Vehicles (DMV) Waymo reported that in its 2017 AV tests its safety drivers disengaged autonomous driving once every 5,500 miles [47]. Waymo reports a disengagement when its evaluation process identifies the event as having 'safety significance', so this rate is almost 100 times worse than the estimated 500,000 miles per accident in 2015 for human drivers. With this poor safety performance, each AV will require a safety driver, making its total cost more than twice today's TNC cost. Of course AV safety will improve over time with more and more testing and R&D but it's anyone's guess as to when an AV will perform as safely as human drivers.

Alternatively, AMoD service can be scaled back to very controlled environments that reduce the risk of accidents by a factor of 100. That direction is also being pursued. For example, Waymo is providing rides to 400 people in the calm, sunny suburb of Chandler, AZ [50]. These AMoD rides use AVs with a safety driver.

One additional piece of evidence also suggests that the cost of AVs is very high. Two proposed contracts show the leasing cost of AV cars and shuttles of well over \$100,000 each per year [51]. EasyMile is charging more than \$27,000 a month per small electric shuttle for cities that sign up for one year of service. Drive.ai charges \$14,000 monthly per vehicle for one year. Considering that a TNC driver (with car) costs \$55,000 per year or \$4,400 per month, it seems unlikely that these are viable business models, except in selective subsidized niche markets.[14]

9. Conclusion

This paper analyzed the impact of three regulations on the ride-hailing app-based platforms or TNCs like Uber and Lyft: (a) a floor under driver wage; (b) a cap on total number of drivers; and (c) a per-trip

[14] According to [51], "Arlington, Texas, a suburb of Dallas, hired Drive.ai to run three on-demand self-driving shuttles in the entertainment district. For the yearlong program, the city will foot 20% of the \$435,000 price tag and a federal grant will cover the rest."

congestion tax that goes to the public transit. We constructed a general equilibrium model to predict market responses to the platform's decision on fares and wages, with and without these regulations. We showed that imposing a wage floor increases driver employment, lowers pickup time, decreases ride cost, and attracts more passengers, over a wide range of parameters. Our analysis suggests that a higher minimum wage benefits both drivers and passengers, at the expense of platform profits. On the other hand, a cap on the number of drivers or vehicles hurts drivers, as the platform benefits by hiring cheaper drivers when supply is limited. Variants of our model were analyzed from other perspectives as well, including platform subsidy, platform competition and autonomous vehicles.

Our study advocates a wage floor for TNC drivers. Our simulation shows that increasing driver wage by 23.3% (from \$22.6 to \$27.86 per hour before expense) will increase the number of TNC vehicles by 23.3% (from 5089 to 6276), increase TNC ridership by 24.6% (from 187 to 233 per min), improve the pickup time by 10% (from 5 min to 4.5 min), decrease the travel cost by 3.6% (from \$33.4 to \$32.2 per trip), and reduce the platform rent by 10.5% (from \$76K to \$68K per hour). This indicates that enforcing a minimum wage for drivers benefits both drivers and passengers. Under the wage floor, the platform is motivated hire more drivers to attract more passengers so as to increase the platform sales. As a consequence, more drivers are hired, more passengers are delivered, each driver earns more, and each passenger spends less. The wage floor squeezes the monopoly profit of the platform and improves the efficiency of the system. It thus boosts the TNC economy without costing taxpayer money.

A congestion surcharge will relieve traffic congestion by reducing the number of TNC vehicles. Numerical simulation suggests that a surcharge of \$2.75 per trip in New York will reduce TNC vehicles by 11.9% (from 5089 to 4480) and TNC rides by 17.1% (from 187 to 155 per min). More importantly, the funds raised from this surcharge can be used to subsidize public transit. In New York, it is estimated that the congestion surcharge will yield \$1M per day. The money goes to the Metropolitan Transportation Authority to upgrade the subway system. It can be used to make public transportation cleaner, faster, and safer, so that more residents will commute by transit. Increased public transit ridership will improve the efficiency of the city's transportation system, and reduce the environmental footprint of transportation, which accounts for 28% percent of the total carbon emissions in the US.

Interest in TNC regulation has been driven by concerns about working conditions of TNC drivers and by the deleterious impact on urban transport of TNC growth. This paper deals only with the impact on driver wage and ride fare. There is a debate whether TNC drivers are more like 'independent contractors' or more like employees [52]. This paper contributes to that debate in showing that TNC driver wages can be significantly increased and passenger fares decreased at the cost of lower TNC profits.

Acknowledgments

This research was supported by National Science Foundation EAGER award 1839843.

References

[1] New York City Taxi and Limousine Commission. Notice of promulgation: Statement of basis and purpose of rules. http://www.nyc.gov/html/tlc/downloads/pdf/driver_income_rules_12_04_2018.pdf, December 2018.

[2] J. A. Parrott and M. Reich. An earning standard for New York City app-based drivers: Economic analysis and policy assessment. The New School, Center for New York City Affairs , 2018.

[3] A. J. Hawkins. Lyft will sue to block NYC's driver pay equity law. Verge, Jan 30, 2019.

[4] J. Burns. Judge blocks Lyft's lawsuit against NYC ride-hail rules, title = Forbes, annote = https://www.forbes.com/sites/janetwburns/2019/05/07/nyc-judge-denies-lyft-lawsuit-against-new-ride-hail-rules/, year = May 7, 2019,.

[5] S. Hsieh. Seattle city council passes resolution to consider raising Uber and Lyft rates. *The Stranger*, Apr 2018.

[6] San Francisco County Transportation Authority. The TNC regulatory landscape. https://www.sfcta.org/sites/default/files/2019-03/TNC_regulatory_020218.pdf, December, 2019.

[7] R.B. Collier, V. Dubai, and C. Carter. The regulation of labor platforms: The politics of the Uber economy, 2019. https://brie.berkeley.edu/sites/default/files/reg-of-labor-platforms.pdf.

[8] S. Butler. Uber loses appeal over driver employment rights. The Guardian, December 19 2018.

[9] A.F. Campbell. California is cracking down on the gig economy. Vox, May 30 2019. https://www.vox.com/policy-and-politics/2019/5/30/18642535/california-ab5-misclassify-employees-contractors.

[10] Uber. Form S1 Registration Statement. United States Securities and Exchange Commission, 2019.

[11] J. Tirole. Regulating the disrupters. livemint, January 2019.

[12] M. Snider. Uber, Lyft rides in NYC could cost more under new minimum wage rule. USA Today, December 5 2018.

[13] J. Castiglione, T. Chang, D. Cooper, J. Hobson, W. Logan, E. Young, B. Charlton, C. Wilson, A. Mislove, L. Chen, and Jiang S. TNCs today: a profile of San Francisco transportation network company activity. San Francisco County Transportation Authority, 2016.

[14] J.W. Hall and A. Kreuger. An analysis of the labor market for Uber's driver-partners in the United States. NBER Working Paper No. 22843, November 2016.

[15] B. Schaller. Empty Seats, Full Streets. Fixing Manhattan's Traffic Problem. Schaller Consulting, December 2017.

[16] A. Shapiro. New York City just voted to cap Uber and Lyft vehicles, and that could make rides more expensive. CNBC, August 8 2018.

[17] J. Castiglione, D. Cooper, B. Sana, D. Tischler, T. Chang, G. D. Erhardt, S. Roy, M. Chen, and A. Mucci. TNCs & Congestion. San Francisco County Transportation Authority, 2018.

[18] S. Banerjee, C. Riquelme, and R. Johari. Pricing in ride-share platforms: A queueing-theoretic approach. https://ssrn.com/abstract=2568258, 2015.

[19] G. P. Cachon, K. M. Daniels, and R. Lobel. The role of surge pricing on a service platform with self-scheduling capacity. *Manufacturing & Service Operations Management*, 19(3):368–384, 2017.

[20] J. Bai, K. C. So, C. S. Tang, X. Chen, and H. Wang. Coordinating supply and demand on an on-demand service platform with impatient customers. *Manufacturing & Service Operations Management*, 2018.

[21] T. A. Taylor. On-demand service platforms. *Manufacturing & Service Operations Management*, 20(4):704–720, 2018.

[22] X. Ban, M. Dessouky, J. Pang, and R. Fan. A general equilibrium model for transportation systems with e-hailing services and flow congestion. http://www-bcf.usc.edu/ maged/publications/taximodel.pdf, 2018.

[23] M. Rysman. The economics of two-sided markets. *Journal of Economic Perspectives*, 23(3):125–43, 2009.

[24] J. Rochet and J. Tirole. Two-sided markets: a progress report. *The RAND Journal of Economics*, 37(3):645–667, 2006.

[25] E. G. Weyl. A price theory of multi-sided platforms. *American Economic Review*, 100(4):1642–72, 2010.

[26] I. Gurvich, M. Lariviere, and A. Moreno. Operations in the on-demand economy: Staffing services with self-scheduling capacity. https://ssrn.com/abstract=2336514, 2016.

[27] L. Zha, Y. Yin, and H. Yang. Economic analysis of ride-sourcing markets. *Transportation Research Part C: Emerging Technologies*, 71:249–266, 2016.

[28] H. Mohring, J. Schroeter, and P. Wiboonchutikula. The values of waiting time, travel time, and a seat on a bus. *The RAND Journal of Economics*, pages 40–56, 1987.

[29] J. Ke, F. Xiao, H. Yang, and J. Ye. Optimizing online matching for ride-sourcing services with multi-agent deep reinforcement learning. *arXiv:1902.06228*, 2019.

[30] Uber technologies, form s-1 registration statement, 2019. https://www.sec.gov/Archives/edgar/data/1543151/000119312519103850/d647752ds1.htm.

[31] Lyft. Form S1 Registration Statement, 2019.

[32] J. P. Schwieterman and M. Livingston. Uber economics: Evaluating the monetary and nonmonetary tradeoffs of TNC and transit services in Chicago, Illinois. *Chaddick Institute Policy Series*, 2018.

[33] D. A. Quarmby. Choice of travel mode for the journey to work: some findings. *Journal of Transport Economics and Policy*, pages 273–314, 1967.

[34] David Card and Alan B Krueger. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *The American Economic Review*, 84(4):772–793, 1994.

[35] David Card and Alan B Krueger. Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply. *American Economic Review*, 90(5):1397–1420, 2000.

[36] David Neumark and William Wascher. Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania: Comment. *American Economic Review*, 90(5):1362–1396, 2000.

[37] Dale Belman and Paul J Wolfson. *What does the minimum wage do?* WE Upjohn Institute, 2014.

[38] John Schmitt et al. Why does the minimum wage have no discernible effect on employment. *Center for Economic and Policy Research*, 22:1–28, 2013.

[39] David Card and Alan B Krueger. *Myth and Measurement: The New Economics of the Minimum Wage-Twentieth-Anniversary Edition*. Princeton University Press, 2015.

[40] D Neumark and W. Wascher. Minimum wages and employment: A review of evidence from the new minimum wage research. NBER Working Paper No. 12663, January 2007.

[41] Campbell McConnell, Stanley Brue, and David Macpherson. *Contemporary labor economics*. McGraw-Hill Education, 2016.

[42] Derek Laing. *Labor Economics*. WW Norton & Company, 2011.

[43] William M. Boal and Michael R. Ransom. Monopsony in the labor market. *Journal of Economic Literature*, 35(1):86–112, 1997.

[44] Avner Shaked and John Sutton. Relaxing price competition through product differentiation. *The Review of Economic Studies*, pages 3–13, 1982.

[45] Avner Shaked and John Sutton. Natural oligopolies. *Econometrica: Journal of the Econometric Society*, pages 1469–1483, 1983.

[46] A. Talebpour and H. S. Mahmassani. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71:143–163, 2016.

[47] Waymo. Waymo Safety Report. https://waymo.com/safetyreport/, 2018.

[48] K. Spieser, K. Treleaven, R. Zhang, E. Frazzoli, D. Morton, and M. Pavone. Toward a systematic approach to the design and evaluation of automated mobility-on-demand systems: A case study in Singapore. In *Road Vehicle Automation*, pages 229–245. Springer, 2014.

[49] F. Rossi, R. Zhang, Y. Hindy, and M. Pavone. Routing autonomous vehicles in congested transportation networks: Structural properties and coordination algorithms. *Autonomous Robots*, 42:1427–1442, 2018.

[50] Alexandria Sage. Waymo unveils self-driving taxi service in arizona for paying customers. *Rueters*, December 2018.

[51] Kaveh Waddell. The true cost of autonomous cars. AXIOS, Dec 9 1918.

[52] J. Tirole. Regulating the disrupters. Project Syndicate, January 2019.

Appendix

A: Proof of Proposition 1

Proof. We prove the result in two steps. (i) First, we consider the case of a passenger located at the origin and $N_I = N - \lambda/\mu$ idle vehicles uniformly and independently distributed in a disk of radius R centered at the origin. We show that the expected distance of the passenger to the closest idle vehicle is $\sqrt{N_{I_0}}\mathbb{E}\{d(N_{I_0})\}\frac{1}{\sqrt{N_I}}(1+\mathcal{O}(\max\{N_{I_0}^{-1}, N_I^{-1}\}))$. (ii) Second, based on the result of part (i), we show that for a city with any two-dimensional area $\mathcal{A}$, the expected shortest distance to an idle vehicle of a passenger is also given by $\sqrt{N_{I_0}}\mathbb{E}\{d(N_{I_0})\}\frac{1}{\sqrt{N_I}}(1+\mathcal{O}(\max\{N_{I_0}^{-1}, N_I^{-1}\}))$.

(i) To prove the first step, let $d(n) := \min(|x_1|, \cdots, |x_n|)$ be the shortest distance to the origin among n idle vehicles where $x_i \in R^2$ is the location of the ith idle vehicle. Then the cumulative distribution function (cdf) of $d(n)$ is

$$\mathbb{P}\{d(n) \le r\} = 1 - \mathbb{P}\{d(n) > r\} = 1 - \mathbb{P}\{|x_i| > r, \ \forall i\} = 1 - (1 - \frac{\pi r^2}{\pi R^2})^n$$

Therefore, the probability density function (pdf) of d is

$$f_{d(n)}(r) = n(1 - \frac{r^2}{R^2})^{n-1}\frac{2r}{R^2}$$

Consequently,

$$
\begin{aligned}
\mathbb{E}\{d(n)\} &= \int_0^R f_{d(n)}(r) \times r \times dr = \int_0^R 2n(1-\frac{r^2}{R^2})^{n-1}\frac{r^2}{R^2}dr \\
&= \frac{2}{3}n(1-\frac{r^2}{R^2})^{n-1}\frac{r^3}{R^2}\Big|_0^R + \int_0^R \frac{2^2}{3}n(n-1)(1-\frac{r^2}{R^2})^{n-2}\frac{r^4}{R^4}dr \\
&= \int_0^R \frac{2^2}{3}n(n-1)(1-\frac{r^2}{R^2})^{n-2}\frac{r^4}{R^4}dr \\
\cdots \\
&= 2^n \frac{\prod_{i=0}^{n-1}(n-i)}{\prod_{i=1}^{n}(2i+1)}\frac{r^{2n+1}}{R^{2n}}\Big|_0^R \\
&= \frac{(2^n n!)^2}{(2n+1)!}R
\end{aligned}
$$

in which the fifth equality follows from an iterative application of integration by parts (similar to the one that leads to the third equality). Therefore, (for any N_{I_0})

$$
\mathbb{E}\{d(N_I)\} = \begin{cases} \mathbb{E}\{d(N_{I_0})\}\prod_{i=N_{I_0}}^{N_I-1}\left(\frac{2i}{2i+1}\right) & \text{if } N_I > N_{I_0} \\ \mathbb{E}\{d(N_{I_0})\}\frac{1}{\prod_{i=N_I}^{N_{I_0}-1}\left(\frac{2i}{2i+1}\right)} & \text{if } N_I < N_{I_0} \end{cases}
$$

Next, we prove that $\prod_{i=m}^{n}\left(\frac{2i}{2i+1}\right) = \frac{\sqrt{m}}{\sqrt{n}}(1+\mathcal{O}(m))$ for $m < n$. We have,

$$
\ln\left(\prod_{i=m}^{n}\left(\frac{2i}{2i+1}\right)\right) = \sum_{i=m}^{n}\ln\left(1-\frac{1}{2i+1}\right).
$$

For $x \in [0, \sim 0.43]$, we have $-x-x^2 \leq \ln(1-x) \leq -x$. Thus,

$$
-\frac{1}{2}\sum_{i=m}^{n}\frac{1}{i+\frac{1}{2}} - \sum_{i=m}^{n}\frac{1}{(2i+1)^2} \leq \sum_{i=m}^{n}\ln\left(1-\frac{1}{2i+1}\right) \leq -\sum_{i=m}^{n}\frac{1}{2i+1}
$$

$$
\Longleftrightarrow
$$

$$
-\frac{1}{2}\sum_{i=m}^{n}\frac{1}{i+\frac{1}{2}} - \frac{1}{2}\sum_{i=m}^{n}\left[\frac{1}{i-\frac{1}{2}} - \frac{1}{i+\frac{1}{2}}\right] \leq \sum_{i=m}^{n}\ln\left(1-\frac{1}{2i+1}\right) \leq -\frac{1}{2}\sum_{i=m}^{n}\frac{1}{2i+1}
$$

$$
\Longrightarrow
$$

$$
-\frac{1}{2}\ln\left(\frac{n}{m}\right) + \mathcal{O}(m^{-1}) - \frac{1}{2}\left[\frac{1}{m-\frac{1}{2}} - \frac{1}{n+\frac{1}{2}}\right] \leq \sum_{i=m}^{n}\ln\left(1-\frac{1}{2i+1}\right) \leq -\frac{1}{2}\ln\left(\frac{n}{m}\right) + \mathcal{O}(m^{-1}).
$$

Therefore, $\prod_{i=m}^{n}\left(\frac{2i}{2i+1}\right) = \frac{\sqrt{m}}{\sqrt{n}}(1+\mathcal{O}(m^{-1}))$ for $m < n$. Consequently,

$$
\begin{aligned}
\mathbb{E}\{d(N_I)\} &= \begin{cases} \mathbb{E}\{d(N_{I_0})\}\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}(1+\mathcal{O}(N_{I_0}^{-1})) & \text{if } N_I > N_{I_0} \\ \mathbb{E}\{d(N_{I_0})\}\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} + (1+\mathcal{O}(N_I^{-1})) & \text{if } N_I < N_{I_0} \end{cases} \\
&= \sqrt{N_{I_0}}\mathbb{E}\{d(N_{I_0})\}\frac{1}{\sqrt{N_I}}(1+\mathcal{O}(\max\{N_{I_0}^{-1}, N_I^{-1}\}))
\end{aligned}
$$

(ii) We cannot directly apply the result of part (i) for a passenger with an arbitrary location inside $\mathcal{A}$. Nevertheless, based on the result of part (i) we can iteratively provide an approximation. For every point $x \in \mathcal{A}$ consider a disk $B_r(x)$ of radius $r > 0$ around x, $r > 0$. We partition $\mathcal{A}$ into two sets $\mathcal{C}$ and $\mathcal{D}$, where (a) $\mathcal{C} = \{x \in \mathcal{A} : B_r(x) \subset \mathcal{A}\}$ is the set of points such that $B_r(x)$ is contained in $\mathcal{A}$, and (b) let $\mathcal{D} := \{x \in \mathcal{A} : B_r(x) \not\subset \mathcal{A}\}$ is the set of points such that $B_r(x)$ is not completely contained in $\mathcal{A}$.

In the following, we first determine an approximate of $\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\}$ for $x \in \mathcal{C}$. We then provide an upper bound approximation for $\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\}$ for $x \in \mathcal{D}$. Putting the results of (ii-a) and (ii-b) together we provide an approximate equation for a general shape $\mathcal{A}$; see (Iteration 1). We then use the approximation provided in (Iteration 1), to provide a better approximation for $\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\}$ for $x \in \mathcal{D}$, and improve our approximation for a general shape $\mathcal{A}$; see (Iteration 2). We repeat the above process iteratively, and determine the best approximation by analyzing the limit point of the above iterative process which leads to the final approximation $\sqrt{N_{I_0}}\mathbb{E}\{d(N_{I_0})\}\frac{1}{\sqrt{N_I}}(1 + \mathcal{O}(\max\{N_{I_0}^{-1}, N_I^{-1}\}))$.

Let $|\mathcal{A}|$ and L denote the area and length of the (assumed smooth) boundary of $\mathcal{A}$, respectively. Then the areas of $\mathcal{C}$ and D are $|\mathcal{C}| := A - \mathcal{O}(Lr)$ and $|\mathcal{D}| := \mathcal{O}(Lr)$, respectively.

(ii-a) First consider an arbitrary point $x \in \mathcal{C}$. Suppose N_I idle vehicles are uniformly and independently distributed in $\mathcal{A}$. Then the pdf of the number m of these vehicles that lie in $B_r(x)$ is a binomial distribution $B(N_I, \frac{\pi r^2}{\mathcal{A}})$ with mean $N_I\frac{\pi r^2}{\mathcal{A}}$ and variance $N_I\frac{\pi r^2}{\mathcal{A}}\left(1 - \frac{\pi r^2}{\mathcal{A}}\right)$.

Let $l := \{\max|y - z| : x, y \in \mathcal{A}\}$ denote the largest distance between two points in $\mathcal{A}$. Then, for every realization of $m > 1$, the conditional expected shortest distance of an idle vehicle to x is given by $\sqrt{N_{I_0}\pi r^2}\mathbb{E}\{d(N_{I_0}\pi r^2)\}\frac{1}{\sqrt{m}}\left(1 + \mathcal{O}\left(\max\{(N_{I_0}\pi r^2)^{-1}, m^{-1}\}\right)\right)\mathbf{1}\{m > 0\} + \mathcal{O}(l)\mathbf{1}\{m = 0\}$ from part (i) where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Note that we modify the expression from part (i) since the shortest distance of an idle vehicle cannot exceed l when $m = 0$. Moreover, we substitute N_{I_0} with $N_{I_0}\frac{\pi r^2}{\mathcal{A}}$ to reflect the fact among the N_{I_0} idle vehicles uniformly distributed in $\mathcal{A}$, on average $N_{I_0}\frac{\pi r^2}{\mathcal{A}}$ vehicles are inside $B_r(x)$.

Taking the expectation with respect to m, the expected shortest distance of an idle vehicle to x is given by

$$\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\}=\mathbb{E}_m\left\{\sqrt{N_{I_0}\frac{\pi r^2}{\mathcal{A}}}\mathbb{E}\{d(N_{I_0}\frac{\pi r^2}{\mathcal{A}})\}\frac{1}{\sqrt{m}}\left(1+\mathcal{O}\left(\max\{(N_{I_0}\frac{\pi r^2}{\mathcal{A}})^{-1},m^{-1}\}\right)\right)\mathbf{1}\{m>0\}+\mathcal{O}(l)\mathbf{1}\{m=0\}\right\} \quad (28)$$

We can write the first term in this expectation as

$$\mathbb{E}_m\left\{\sqrt{N_{I_0}\frac{\pi r^2}{\mathcal{A}}}\mathbb{E}\{d(N_{I_0}\frac{\pi r^2}{\mathcal{A}})\}\frac{1}{\sqrt{m}}\left(1+\mathcal{O}\left(\max\{(N_{I_0}\frac{\pi r^2}{\mathcal{A}})^{-1}, m^{-1}\}\right)\right)\mathbf{1}\{m>0\}\right\} =$$

$$\sqrt{N_{I_0}\frac{\pi r^2}{\mathcal{A}}}\mathbb{E}\{d(N_{I_0}\frac{\pi r^2}{\mathcal{A}})\}\frac{1}{\sqrt{N_I\frac{\pi r^2}{\mathcal{A}}}}\mathbb{E}_m\left\{\frac{1}{\sqrt{1+\frac{m-N_I\frac{\pi r^2}{\mathcal{A}}}{N_I\frac{\pi r^2}{\mathcal{A}}}}}\left(1+\mathcal{O}\left(\max\{(N_{I_0}\frac{\pi r^2}{\mathcal{A}})^{-1}, (N_I\frac{\pi r^2}{\mathcal{A}})^{-1}(1+\frac{m-N_I\frac{\pi r^2}{\mathcal{A}}}{N_I\frac{\pi r^2}{\mathcal{A}}})^{-1}\}\right)\right)\mathbf{1}\{m>0\}\right\}$$

$$= \sqrt{N_{I_0}}\mathbb{E}\{d(N_{I_0}\frac{\pi r^2}{\mathcal{A}})\}\frac{1}{\sqrt{N_I}}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}, N_I^{-1}\}\right)\right)\mathbb{P}\{m>0\} \quad (29)$$

where the last equality is by Taylor expansion for terms $\frac{1}{\sqrt{1+\frac{m-N_I\frac{\pi r^2}{\mathcal{A}}}{N_I\frac{\pi r^2}{\mathcal{A}}}}}$ and $(1+\frac{m-N_I\frac{\pi r^2}{\mathcal{A}}}{N_I\frac{\pi r^2}{\mathcal{A}}})^{-1}$, along with the fact that $\mathbb{E}_m\{m - N_I\frac{\pi r^2}{\mathcal{A}}\} = 0$ and $\mathbb{E}_m\left\{\left(\frac{m-N_I\frac{\pi r^2}{\mathcal{A}}}{N_I\frac{\pi r^2}{\mathcal{A}}}\right)^2\right\} = \mathcal{O}(N_I^{-1})$.

Moreover, we have $\mathbb{P}\{m=0\} = (1-\frac{\pi r^2}{\mathcal{A}})^{N_I}$. Using inequality $\ln((1-x)^n) \le -nx$, we have

$$\mathbb{P}\{m=0\} = \exp(\ln((1-\frac{\pi r^2}{\mathcal{A}})^{N_I})) \le \exp(-N_I\frac{\pi r^2}{\mathcal{A}}).$$

Set $r = \mathcal{O}(N_I^{\frac{-1}{2}+\delta})$ for some $\delta \in (0, 1/2]$. Then we have

$$\mathbb{P}\{m=0\} = \mathcal{O}(\exp(-N_I^{2\delta})). \tag{30}$$

Consequently, $\mathbb{P}\{m>0\} = 1-\mathbb{P}\{m=0\} = 1+\mathcal{O}(\exp(-N_I^{2\delta}))$. Substituting $\mathbb{P}\{m>0\}$ and $\mathbb{P}\{m=0\}$ in (28), we get

$$\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\} = \mathbb{E}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, \exp(-N_I^{2\delta})\}\right)\right) \tag{31}$$

for $x \in \mathcal{C}(x)$ and $\delta \in (0, 1/2]$.

(ii-b) For points $x \in \mathcal{D}$, consider the intersection of $B_r(x)$ and $\mathcal{A}$. We assume that $\frac{|B_r(x)\cap\mathcal{A}|}{r^2} = \mathcal{O}(1)$ since $\mathcal{A}$ has a smooth boundary. (Here $|\mathcal{R}|$ is the area of $\mathcal{R}$.) Apply an argument similar to the one given in (ii-a) for $B_r(x)\cap\mathcal{A}$, and we can show that with large enough probability the closest idle vehicles to x has a distance smaller than r. More precisely,

$$\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\} \le \mathcal{O}(r) + \mathcal{O}(exp(-N_I r^2)) = \mathcal{O}(N_I^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_I^{2\delta})) \tag{32}$$

for $x \in \mathcal{D}$, where the last equality follows from $r = \mathcal{O}(N_I^{\frac{-1}{2}+\delta})$. Therefore, from (32),

$$\begin{aligned}
\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\} - \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}\{d(x)|N_{I_0}\,\text{idle vehicles}\} &\le \mathcal{O}(N_I^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_I^{2\delta})) - \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}\{d(x)|N_{I_0}\,\text{idle vehicles}\} \\
&= \mathcal{O}(N_I^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_I^{2\delta})) + \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\left[\mathcal{O}(N_{I_0}^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_{I_0}^{2\delta}))\right] \\
&= \mathcal{O}(N_I^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_I^{2\delta})) + \left(\mathcal{O}(N_I^{\frac{-1}{2}}N_{I_0}^{\delta}) + \mathcal{O}(N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\right)
\end{aligned}$$

for $x \in \mathcal{D}$, where the first equality follows from

$$0 \le \left|\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\right| \le \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\left[\mathcal{O}(N_{I_0}^{\frac{-1}{2}+\delta}) + \mathcal{O}(exp(-N_{I_0}^{2\delta}))\right].$$

Therefore, for $x \in \mathcal{D}$, we can write

$$\mathbb{E}\{d(x)|N_I\,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}\{d(x)|N_{I_0}\,\text{idle vehicles}\} + \mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right) \tag{33}$$

(Iteration 1) Using (31) for $x \in \mathcal{C}$ and (33) for $x \in \mathcal{D}$, along with $\mathbb{P}\{x \in \mathcal{D}\} = \frac{|\mathcal{D}|}{|\mathcal{A}|} = \mathcal{O}(\frac{Lr}{|\mathcal{A}|}) = \mathcal{O}(N_I^{-\frac{1}{2}+\delta})$, we have,

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I\,\text{idle vehicles}\} = \Big[\mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{C}\}|N_I\,\text{idle vehicles}\}\Big] + \Big[\mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{D}\}|N_I\,\text{idle vehicles}\}\Big]$$

$$= \left[\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{C}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\left(1+\mathcal{O}\Big(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta})\}\Big)\right)\right](1-\mathbb{P}\{x\in\mathcal{D}\})$$

$$+\left[\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{D}}\{d(x)|N_{I_0}\,\text{idle vehicles}\} + \mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right)\right]\mathbb{P}\{x\in\mathcal{D}\}$$

$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}$$

$$+\left[\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{C}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\mathcal{O}\Big(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta})\}\Big)(1-\mathbb{P}\{x\in\mathcal{D}\})\right]$$

$$+\left[\mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right)\mathbb{P}\{x\in\mathcal{D}\}\right]$$

$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}$$

$$+\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\Bigg[\frac{\mathbb{E}_{x\in\mathcal{C}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}}{\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}}\mathcal{O}\Big(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta})\}\Big)(1-\mathbb{P}\{x\in\mathcal{D}\})$$

$$+\frac{\sqrt{N_I}}{\sqrt{N_{I_0}}}\frac{1}{\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}}\mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right)\mathbb{P}\{x\in\mathcal{D}\}\Bigg]$$

$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}$$

$$+\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\Bigg[\mathcal{O}(1)\mathcal{O}\Big(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta})\}\Big)(1-\mathbb{P}\{x\in\mathcal{D}\})$$

$$+\frac{\sqrt{N_I}}{\sqrt{N_{I_0}}}\frac{1}{\mathcal{O}(N_{I_0}^{\frac{-1}{2}})}\mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right)\mathbb{P}\{x\in\mathcal{D}\}\Bigg]$$

$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}$$

$$+\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\Bigg[\mathcal{O}\Big(\max\{N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta})\}\Big)(1-\mathcal{O}(N_I^{\frac{-1}{2}+\delta}))$$

$$+\sqrt{N_I}\mathcal{O}\left(\max\{N_I^{\frac{-1}{2}+\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}}N_{I_0}^{\delta}, N_{I_0}^{\frac{1}{2}}N_I^{-\frac{1}{2}}exp(-N_{I_0}^{2\delta})\}\right)\mathcal{O}(N_I^{\frac{-1}{2}+\delta})\Bigg]$$

$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_{I_0}\,\text{idle vehicles}\}\Bigg(1+\mathcal{O}(\max N_{I_0}^{-1}N_I^{1-2\delta}, N_I^{-2\delta}, exp(-N_I^{2\delta}), N_I^{\frac{-1}{2}+2\delta}, N_I^{\delta}exp(-N_I^{2\delta}), N_I^{-\frac{1}{2}+\delta}N_{I_0}^{\delta},$$

$$N_{I_0}^{\frac{1}{2}}N_I^{\frac{-1}{2}+\delta}exp(-N_{I_0}^{2\delta}))\Bigg),$$

for $\delta\in(0,\frac{1}{2}]$.

Setting $\delta=\frac{1}{8}$ we have,

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I\,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}N_I^{\frac{3}{4}}, N_I^{-\frac{1}{4}}, N_I^{-\frac{3}{8}}N_{I_0}^{\frac{1}{8}}\}\right)\right). \quad (34)$$

where we neglect $\mathcal{O}(exp(-N_I^{\frac{1}{4}}))$, $\mathcal{O}(N_I^{0.125}exp(-N_I^{\frac{1}{4}}))$ and $\mathcal{O}(N_{I_0}^{0.5}N_I^{-0.375}exp(-N_I^{\frac{1}{4}}))$ in comparison to $\mathcal{O}\left(N_I^{-\frac{1}{4}}\right)$.

(Iteration 2) We can now use (34) to provide a better approximation for $x\in\mathcal{D}$ in part (ii-b) and iterate the process. Note that the number of idle vehicles in $\mathcal{D}$ (on average) is equal to $\frac{|\mathcal{D}|}{|\mathcal{A}|}N_I=\mathcal{O}(r)N_I=N_I^{\frac{1}{2}+\delta}$;

similarly, we need to substitute N_{I_0} with $N_{I_0}\mathcal{O}(r) = N_{I_0} N_I^{\frac{-1}{2}+\delta}$ in (34).[15] Therefore,

$$\mathbb{E}\{d(x)|N_I \,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{\frac{7}{8}-\frac{1}{4}\delta}, N_I^{-\frac{1}{8}-\frac{1}{4}\delta}, N_{I_0}^{\frac{1}{8}} N_I^{\frac{-1}{4}-\frac{1}{4}\delta}\}\right)\right) \tag{35}$$

for $x \in \mathcal{D}$.[16]

Using (31) for $x \in \mathcal{C}$ and (35) for $x \in \mathcal{D}$, along with $|\mathcal{D}| = \mathcal{O}(r) = \mathcal{O}(N^{\frac{-1}{2}+\delta})$, we have

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I \,\text{idle vehicles}\} = \mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{C}\}|N_I\,\text{idle vehicles}\} + \mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{D}\}|N_I\,\text{idle vehicles}\}$$
$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{1-2\delta}, N_I^{-2\delta}, N_{I_0}^{-1} N_I^{\frac{3}{8}+\frac{3}{4}\delta}, N_I^{-\frac{5}{8}+\frac{3}{4}\delta}, N_{I_0}^{\frac{1}{8}} N_I^{-\frac{3}{4}+\frac{3}{4}\delta}, exp(-N_I^{2\delta})\}\right)\right).$$

Setting $\delta = \frac{5}{22}$, we have an improved approximation

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I \,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{\frac{6}{11}}, N_I^{-\frac{5}{11}}, N_{I_0}^{\frac{1}{8}} N_I^{-\frac{5}{11}-\frac{1}{8}}\}\right)\right), \tag{36}$$

where we neglect $\mathcal{O}(exp(-N_I^{\frac{5}{11}}))$ in comparison to $\mathcal{O}\left(N_I^{-\frac{5}{11}}\right)$.

(Iteration K) We can iterate the same process similar to the one described in iteration 2. Assume that at the end of iteration $K-1$ we show that

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I \,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{\alpha(K-1)}, N_I^{-(1-\alpha(K-1))}, N_{I_0}^{\frac{1}{8}} N_I^{-(1-\alpha(K-1))-\frac{1}{8}}\}\right)\right). \tag{37}$$

where $\alpha(K-1) \in (0, \frac{1}{2}]$. Note that in iteration 2 we have $\alpha(2) = \frac{6}{11}$. We now use (37) to provide a better approximation for $x \in \mathcal{D}$ and iterate the process another time. Once again note that the number of idle vehicles in $\mathcal{D}$ (on average) is equal to $\frac{|\mathcal{D}|}{|\mathcal{A}|} N_I = \mathcal{O}(r) N_I = N_I^{\frac{1}{2}+\delta}$; similarly, we substitute N_{I_0} with $N_{I_0}\mathcal{O}(r) = N_{I_0} N_I^{\frac{-1}{2}+\delta}$ in (36). Consequently,

$$\mathbb{E}\{d(x)|N_I \,\text{idle vehicles}\} =$$
$$\frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{\frac{1+\alpha(K-1)}{2}-(1-\alpha(K-1))\delta}, N_I^{-\frac{1-\alpha(K-1)}{2}-(1-\alpha(K-1))\delta}, N_{I_0}^{\frac{1}{8}} N_I^{-\frac{1-\alpha(K-1)}{2}-(1-\alpha(K-1))\delta-\frac{1}{8}}\}\right)\right), \tag{38}$$

for $x \in \mathcal{D}$.

Using (31) for $x \in \mathcal{C}$ and (38) for $x \in \mathcal{D}$, along with $|\mathcal{D}| = \mathcal{O}(r) = \mathcal{O}(N^{\frac{-1}{2}+\delta})$, we have

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I \,\text{idle vehicles}\} = \mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{C}\}|N_I\,\text{idle vehicles}\} + \mathbb{E}_x\{d(x)\mathbf{1}\{x\in\mathcal{D}\}|N_I\,\text{idle vehicles}\}$$
$$= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}} \mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\} \left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1} N_I^{1-2\delta}, N_I^{-2\delta}, N_{I_0}^{-1} N_I^{\frac{\alpha(K-1)}{2}+\alpha(K-1)\delta}, N_I^{-\frac{2-\alpha(K-1)}{2}+\alpha(K-1)\delta}, N_{I_0}^{\frac{1}{8}} N_I^{-\frac{2-\alpha(K-1)}{2}+\alpha(K-1)\delta-\frac{1}{8}}, exp(-N_I^{2\delta})\}\right)\right).$$

[15] We note that the number of vehicles in $\mathcal{D}$ is a random variable. Nevertheless, considering the variation in the number of vehicles around the average of $N_I^{\frac{1}{2}+\delta}$ one can follow a similar approach we used in (ii-a) using Taylor expansion, and show that the error term due to such a randomness is smaller than other error terms and does not affect the result.

[16] We note that the closest idle vehicles for $x \in \mathcal{D}$ does not necessarily belong to $\mathcal{D}$ and can be inside $\mathcal{C}$. Nevertheless, we can still use (34) for the expected distance of the closest idle vehicle for x. This is because for area $\mathcal{D}$ we can follow an argument similar to the one that leads to (34), and divide it to an interior region $\tilde{\mathcal{C}}$ and an exterior region $\tilde{\mathcal{D}}$, ; however, we consider the border of $\mathcal{D}$ that separates it from $\mathcal{C}$ as a part of the interior region $\tilde{\mathcal{C}}$ (and not $\tilde{\mathcal{D}}$). Consequently, equation (34) is also applicable for $x \in \mathcal{D}$.

Setting $\delta = \frac{2-\alpha(K-1)}{2(2+\alpha(K-1))}$, we have an improved approximation for iteration K as

$$\begin{aligned}\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I\,\text{idle vehicles}\} &= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}N_I^{\frac{2\alpha(K-1)}{2+\alpha(K-1)}}, N_I^{-\frac{2-\alpha(K-1)}{2+\alpha(K-1)}}, N_{I_0}^{\frac{1}{8}}N_I^{-\frac{2-\alpha(K-1)}{2+\alpha(K-1)}-\frac{1}{8}}\}\right)\right)\\ &= \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}N_I^{\alpha(K)}, N_I^{1-\alpha(K)}, N_{I_0}^{\frac{1}{8}}N_I^{1-\alpha(K)-\frac{1}{8}}\}\right)\right), \quad (39)\end{aligned}$$

where $\alpha(K) = \frac{2\alpha(K-1)}{2+\alpha(K-1)}$ and we neglect $\mathcal{O}(\exp(-N_I^{\frac{2-\alpha(K-1)}{2+\alpha(K-1)}}))$ in comparison to $\mathcal{O}\left(N_I^{-\frac{2-\alpha(K-1)}{2+\alpha(K-1)}}\right)$.

Iterating the above process, it is easy to show that the sequence of $\alpha(K) = \frac{2\alpha(K-1)}{2+\alpha(K-1)}$ converges to $\alpha^* = 0$. Therefore, we have

$$\mathbb{E}_{x\in\mathcal{A}}\{d(x)|N_I\,\text{idle vehicles}\} = \frac{\sqrt{N_{I_0}}}{\sqrt{N_I}}\mathbb{E}_x\{d(x)|N_{I_0}\,\text{idle vehicles}\}\left(1+\mathcal{O}\left(\max\{N_{I_0}^{-1}, N_I^{-1}, N_{I_0}^{\frac{1}{8}}N_I^{-\frac{9}{8}}\}\right)\right).$$

□

B: Proof of Proposition 2

Proof. We can represent (N, λ, p_f, p_d) as a function of total travel cost c and driver wage w:

$$\begin{cases} \lambda = \lambda_0\left[1 - F_p(c)\right] & (40a)\\ N = N_0 F_d(w) & (40b)\\ p_f = \dfrac{1}{\beta}\left[c - \alpha t_w\left(N_0F_d(w) - \lambda_0[1-F_p(c)]/\mu\right)\right] & (40c)\\ p_d = \dfrac{wN}{\lambda} = w\dfrac{N_0F_d(w)}{\lambda_0\left[1-F_p(c)\right]} & (40d)\end{cases}$$

where (40c) and (40d) are obtained based on the definition of c and w, i.e., (2) and (6). To prove Proposition 2, it suffices to show that there exists $c > 0$ and $w > 0$ such that

$$\begin{cases} c > \alpha t_w\left(N_0F_d(w) - \lambda_0[1-F_p(c)]/\mu\right) & (41a)\\ F_p(c) < 1 & (41b)\\ F_d(w) > 0 & (41c)\end{cases}$$

Note that (41a) corresponds to $p_f > 0$, and (41b) and (41c) guarantee that $p_d > 0$, $N > 0$ and $\lambda > 0$. As t_w is decreasing (see Assumption 1), it suffices to prove that there exists $c > 0$ and $w > 0$ such that $F_p(c) < 1$, $F_d(w) > 0$ and that:

$$c > \alpha t_w\left(N_0 - \lambda_0[1-F_p(c)]/\mu\right). \quad (42)$$

The left hand side of (42) is an increasing function of c, while the right hand side is decreasing function. Define $c^* = \inf_c\{c|F_p(c) = 1\}$. It suffices to show that $c^* > \alpha t_w(N_0)$. This is equivalent to $F_p(\alpha t_w(N_0)) < 1$, which completes the proof. □

C: Proof of Proposition 3

Proof. To prove Proposition 3, we first show that there is at most one solution to (12). Then we show that this solution exists, and it coincides with the globally optimal solution to (9).

Uniqueness: let $f_p(c)$ and $f_d(w)$ be the probability density function of c and w, respectively. Since c and w are subject to uniform distribution, we have: $f_p(c) = e_p * \mathbb{1}_{e_p c \leq 1}$ and $f_d(w) = e_d * \mathbb{1}_{e_d w \leq 1}$, where $\mathbb{1}_A$ is the indicator function of A. Assume for the moment that (12) admits a non-trivial solution. Denote it as $(\tilde{p}_f, \tilde{p}_d, \tilde{\lambda}, \tilde{N})$, and denote $\tilde{c}$ and $\tilde{w}$ as the corresponding passenger cost and driver wage, respectively. We first show that $e_p\tilde{c} \leq 1$ and $e_d\tilde{w} \leq 1$. This is because if $e_p\tilde{c} > 1$, then $\lambda = N = 0$, this is a trivial solution. If $e_d\tilde{w} > 1$, we can decrease $\tilde{p}_d$ without affecting $\tilde{\lambda}$, indicating that $\dfrac{\partial \Pi}{\partial p_d} < 0$, which contradicts with (12b). Therefore, we can rewrite (12c) and (12d) as:

$$\begin{cases} \lambda = \lambda_0 \left[1 - e_p\left(\dfrac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f\right)\right] \\ N = N_0 e_d \left(\dfrac{\lambda p_d}{N}\right). \end{cases}$$

This can be further simplified to:

$$\lambda = \lambda_0 \left[1 - e_p\left(\frac{\alpha M}{\sqrt{\sqrt{N_0 e_d \lambda p_d} - \lambda/\mu}} + \beta p_f\right)\right] \tag{43}$$

It suffices to show that there exists at most one set of $(\tilde{p}_f, \tilde{p}_d, \tilde{\lambda})$ that satisfies (12a), (12b) and (43). Using the implicit function theorem on (43), we can derive $\dfrac{\partial \lambda}{p_f}$ and $\dfrac{\partial \lambda}{p_d}$, thus (12a)- (12b) becomes:

$$\begin{cases} \dfrac{-\lambda_0 \beta e_p \left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2}}{\left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2} - \dfrac{1}{2}\alpha M e_p \lambda_0 \left[\dfrac{1}{2}\sqrt{N_0 e_d p_d/\lambda} - 1/\mu\right]}(p_f - p_d) + \lambda = 0 & (44a) \\ \dfrac{\dfrac{1}{4}\lambda_0 e_p \alpha M \sqrt{N_0 e_d \lambda / p_d}}{\left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2} - \dfrac{1}{2}\alpha M e_p \lambda_0 \left[\dfrac{1}{2}\sqrt{N_0 e_d p_d/\lambda} - 1/\mu\right]}(p_f - p_d) - \lambda = 0 & (44b) \end{cases}$$

This reduces to:

$$\begin{cases} \dfrac{1}{4}\lambda_0 e_p \alpha M p_f \sqrt{N_0 e_d \lambda / p_d} = \lambda \left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2} + \alpha M e_p \lambda_0 \lambda / 2\mu & (45a) \\ \beta \sqrt{p_d/\lambda}\left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2} = \dfrac{1}{4}\alpha M \sqrt{N_0 e_d} & (45b) \\ p_f \lambda_0 e_p \beta \sqrt{N_0 e_d} = \sqrt{N_0 e_d}\lambda + 2\sqrt{\lambda p_d} e_p \lambda_0 \beta/\mu, & (45c) \end{cases}$$

where (45a) directly follows from (44b), (45b) follows from (44a) and (44b), and (45c) are derived by plugging (45b) into (45a).

Assume there is another solution, denoted $(p'_f, p'_f, \lambda', N')$. Without loss of generality, suppose $p'_f \leq \tilde{p}_f$. If $p'_f < \tilde{p}_f$, then there are three cases:

- $p'_d \geq \tilde{p}_d$. Based on (43), we have $\lambda' > \tilde{\lambda}$. However, (45c) dictates that $\lambda' < \tilde{\lambda}$. A contradiction.

- $p_d' < \tilde{p}_d$ and $\sqrt{N_0 e_d p_d' \lambda'} - \lambda'/\mu \geq \sqrt{N_0 e_d \tilde{p}_d \tilde{\lambda}} - \tilde{\lambda}/\mu$. Note that (45a) is equivalent to:

$$\frac{1}{4}\lambda_0 e_p \alpha M \sqrt{N_0 e_d \lambda} p_f = \sqrt{p_d \lambda}\left[\left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{3/2} + \alpha M e_p \lambda_0 / 2\mu\right] \tag{46}$$

  Therefore, based on (46), $p_d' \lambda' < \tilde{p}_d \tilde{\lambda}$. On the other hand, since $\sqrt{N_0 e_d p_d' \lambda'} - \lambda'/\mu \geq \sqrt{N_0 e_d \tilde{p}_d \tilde{\lambda}} - \tilde{\lambda}/\mu$ and $p_f' < \tilde{p}_f$, based on (43), we have $\lambda' > \tilde{\lambda}$. This indicates that $\sqrt{N_0 e_d p_d' \lambda'} - \lambda'/\mu < \sqrt{N_0 e_d \tilde{p}_d \tilde{\lambda}} - \tilde{\lambda}/\mu$. A contradiction.

- $p_d' < \tilde{p}_d$ and $\sqrt{N_0 e_d p_d' \lambda'} - \lambda'/\mu < \sqrt{N_0 e_d \tilde{p}_d \tilde{\lambda}} - \tilde{\lambda}/\mu$. Note that (45b) is equivalent to:

$$\beta\sqrt{p_d}(\sqrt{N_0 e_d p_d} - \sqrt{\lambda}/\mu)\left(\sqrt{N_0 e_d p_d \lambda} - \lambda/\mu\right)^{1/2} = \frac{1}{4}\alpha M \sqrt{N_0 e_d} \tag{47}$$

  As $p_d' < \tilde{p}_d$ and $\sqrt{N_0 e_d p_d' \lambda'} - \lambda'/\mu < \sqrt{N_0 e_d \tilde{p}_d \tilde{\lambda}} - \tilde{\lambda}/\mu$, (47) indicates that $\lambda' < \tilde{\lambda}$. On the other hand, (45b) also indicates that $p_d'/\lambda' > \tilde{p}_d/\tilde{\lambda}$. Note that (43) can be written as

$$\lambda = \lambda_0\left[1 - e_p\left(\frac{\alpha M}{\sqrt{\sqrt{\lambda} N_0 e_d p_d} - \lambda/\mu} + \beta p_f\right)\right] \tag{48}$$

  As $p_f' < \tilde{p}_f$ and $p_d'/\lambda' > \tilde{p}_d/\tilde{\lambda}$, (48) indicates that $\lambda' > \tilde{\lambda}$. A contradiction.

If $p_f' = \tilde{p}_f$ and $p_d' < \tilde{p}_d$, then we can find a contradiction by exactly the same argument. In addition, if $p_f' = \tilde{p}_f$ and $p_d' = \tilde{p}_d$, then λ and N are uniquely determined by (45).

Existence and Optimality: Based on (40), we can represent (N, λ, p_f, p_d) as a function of total travel cost c and driver wage w. Therefore, the platform profit $\lambda(p_f - p_d)$ is a function of c and w:

$$R = \lambda_0[1 - F_p(c)] * \left[\frac{c}{\beta} - \frac{\alpha}{\beta} t_w\left(N_0 F_d(w) - \lambda_0[1 - F_p(c)]/\mu\right) - w\frac{N_0 F_d(w)}{\lambda_0\left[1 - F_p(c)\right]}\right] \tag{49}$$

Note that R is well-defined if $\lambda > 0$, $N > 0$, and the number of idle vehicle is positive, i.e., $N_0 F_d(w) > \lambda_0[1 - F_d(c)]/\mu$. Therefore, R is a continuous function of c and w defined in

$$\mathcal{D} = \{(c, w) | 0 \leq c < 1/e_r, 0 < w \leq 1/e_r, N_0 e_d w > \lambda_0(1 - e_p c)/\mu\}.$$

It can be verified that $R \leq 0$ on the boundary of $\mathcal{D}$. Based on our assumption, there exists $(c^*, w^*) \in \mathcal{D}$ at which $R > 0$. This indicates that the optimal solution to (9) is in the interior of $\mathcal{D}$, which solves (12). Since (12) has at most one solution, this completes the proof.

□

D: Proof of Theorem 1

Proof. We will first prescribe a procedure to compute the optimal solution to (15), then we use the procedure to prove that $\nabla_+ N^*(\tilde{w}) > 0$ and $\nabla_+ \lambda^*(\tilde{w}) > 0$.

Computation procedure: Let $(p_f^*, p_d^*, \lambda^*, N^*)$ denote the optimal solution to (15). Note that (16b) at equality uniquely defines an increasing mapping from N to λp_d. Denote this mapping as $\lambda p_d = g(N)$, then

(16b) reduces to $\lambda p_d \geq g(N)$, (16c) reduces to $\lambda p_d \geq Nw$. Therefore, (16b) and (16c) can be combined as $\lambda p_d \geq \max\{g(N), wN\}$, and (15) is equivalent to:

$$\max_{p_f, N} \quad \lambda p_f - \max\{g(N), wN\} \tag{50}$$
$$\text{s.t. } \lambda = \lambda_0 \left[1 - F_p\left(\frac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f\right)\right].$$

Consider the following problem:

$$\max_{p_f, N} \quad \lambda p_f - g(N) \tag{51}$$
$$\text{s.t. } \lambda = \lambda_0 \left[1 - F_p\left(\frac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f\right)\right]. \tag{52}$$

Let $(\bar{p_f}, \bar{N}, \bar{\lambda})$ be the optimal solution to (51), and denote the optimal value as $\bar{R}$. To facilitate our discussion, define $\Pi_1(N)$ as a function representing the optimal value of (51) for any *given* N, then $\bar{R} = \max_N \Pi_1(N)$. Similarly, define:

$$\max_{p_f, N} \quad \lambda p_f - wN \tag{53}$$
$$\text{s.t. } \lambda = \lambda_0 \left[1 - F_p\left(\frac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f\right)\right]. \tag{54}$$

Let $(\hat{p_f}, \hat{N}, \hat{\lambda})$ be the optimal solution to (53), and denote the optimal value as $\hat{R}$. Define $\Pi_2(N)$ as the optimal value of (53) for any given N, then $\hat{R} = \max_N \Pi_2(N)$.

Lemma 1. *There are three cases for the solution to (15):*
(a) if $\Pi_1(\bar{N}) < \Pi_2(\bar{N})$, then the solution to (15) is given by that to (51);
(b) if $\Pi_2(\hat{N}) < \Pi_1(\hat{N})$, then the solution to (15) is given by that to (53);
(c) if $\Pi_1(\bar{N}) \geq \Pi_2(\bar{N})$ and $\Pi_2(\hat{N}) \geq \Pi_1(\hat{N})$, then the solution to (15) is given by:

$$\max_{p_f, N} \quad \lambda p_f - wN \tag{55}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\left(\frac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f\right)\right] & \text{(56a)} \\ N = N_0 F_d(w) & \text{(56b)} \end{cases}$$

The proof of Lemma 1 can be found in Appendix F. It provides a procedure to compute the solution to (15): first compute $\bar{N}$ and $\hat{N}$ by solving (51) and (53) respectively, then identify which case of (a), (b), (c) applies, and obtain N^* correspondingly.

Driver: We first prove that $\nabla_+ N^*(\tilde{w}) > 0$. Based on Lemma 1, there are three cases: (a) $\Pi_1(\bar{N}) < \Pi_2(\bar{N})$, (b) $\Pi_2(\hat{N}) < \Pi_1(\hat{N})$, (c) $\Pi_1(\bar{N}) \geq \Pi_2(\bar{N})$ and $\Pi_2(\hat{N}) \geq \Pi_1(\hat{N})$. In case (a), the constraint (16c) is inactive at the optimal solution to (15) (see proof of Lemma 1). Therefore, case (a) corresponds to $w \leq \tilde{w}$. When $w > \tilde{w}$, the minimum wage constraint is active, and either case (b) and (c) applies.

Note that in case (c), both (16b) and (16c) are active at the optimal solution (see proof of Lemma 1). Therefore, when w increases, N also increases. Then it suffices to show that there exists $\bar{w} > \tilde{w}$, such that the condition of case (c) holds when $\tilde{w} < w \leq \bar{w}$.

Assume not, then there exists $\bar{w}' > \tilde{w}$ such that the conditions of case (b) holds when $\tilde{w} \leq w \leq \bar{w}'$. In this case, we have $\Pi_1(\bar{N}) < \Pi_2(\bar{N})$ for $w \leq \tilde{w}$ and $\Pi_2(\hat{N}) < \Pi_1(\hat{N})$ for $\tilde{w} \leq w \leq \bar{w}'$. Since $\Pi_2(N)$ is continuous with respect to w, let w approaches $\tilde{w}$ from the let, then we have $\Pi_1(\tilde{N}) \leq \Pi_2(\tilde{N})$. Let w approaches $\tilde{w}$ from the right, then we have $\Pi_1(\tilde{N}) \leq \Pi_2(\tilde{N})$. This implies that $\Pi_1(\tilde{N}) = \Pi_2(\tilde{N})$.

In other words, the optimal solution to (51) and (53) are the same when $w = \tilde{w}$. This indicates that partial derivative of their objective functions with respect to N at $N = N^*$ are the same, i.e., $\nabla g(N^*) = \tilde{w}$. Note that $w = \tilde{w}$ indicates $N^* = \tilde{N}$. Therefore, we have $\nabla g(\tilde{N}) = \tilde{w}$. Since $x = g(N)$ is defined as $N = N_0 F_d(x/N)$, applying implicit function theorem, we can obtain $\nabla g(N)$ and derive that $\nabla g(\tilde{N}) = \tilde{w}$ is equivalent to

$$\frac{1 + f_d(\tilde{\lambda}\tilde{p}_2/\tilde{N}) N_0 \tilde{\lambda}\tilde{p}_2/\tilde{N}^2}{f_d(\tilde{\lambda}\tilde{p}_2/\tilde{N}) N_0/\tilde{N}} = \frac{\tilde{\lambda}\tilde{p}_2}{\tilde{N}}, \tag{57}$$

which is clearly impossible. A contradiction.

Passenger: We now prove that $\nabla_+ \lambda^*(\tilde{w}) > 0$. Consider w such that $\tilde{w} < w < \bar{w}$. In this regime, both (16b) and (16c) are active at the optimal solution. Therefore (15) is equivalent to:

$$\max_{p_f \geq 0} \lambda p_f - Nw \tag{58}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p\Big(\alpha t_w(N - \lambda/\mu) + \beta p_f\Big)\right] & \text{(59a)} \\ N = N_0 F_d(w) & \text{(59b)} \end{cases}$$

We can plug (40a), (40b) and (40c) into the objective function of (58), transforming (58) to:

$$\max_{c} \quad \lambda_0[1 - F_p(c)] \cdot \frac{1}{\beta}\left[c - \alpha t_w\Big(N_0 F_d(w) - \lambda_0[1 - F_p(c)]/\mu\Big)\right] - N_0 F_d(w) \cdot w. \tag{60}$$

Note that w is exogenous. The first order condition dictates that the derivative of (60) with respect to c is 0 at the optimal solution:

$$\Phi(c, w) = -\lambda_0 f_p(c) \cdot \frac{1}{\beta}\left[c - \alpha t_w(N_I)\right] + \lambda_0[1 - F_p(c)] \cdot \frac{1}{\beta}\left[1 - \alpha t_w'(N_I)\lambda_0 f_p(c)/\mu\right] = 0, \tag{61}$$

where $N_I = N_0 F_d(w) - \lambda_0[1 - F_p(c)]/\mu$. Based on (40a), λ is a decreasing function of c, thus it suffices to show that the positive partial derivative of c with respect to w is strictly negative, i.e., $\frac{\partial c}{\partial_+ w} < 0$. According to implicit function theorem, we have:

$$\frac{\partial c}{\partial_+ w} = -\frac{\partial \Phi}{\partial_+ w} \Big/ \frac{\partial \Phi}{\partial c}.$$

If c^* is local maximum for (60), then $\Phi(c, w) > 0$ in the neighborhood $c^* - \epsilon < c < c^*$ and $\Phi(c, w) < 0$ in $c^* < c < c^* + \epsilon$. This indicates that $\frac{\partial \Phi}{\partial c^*} < 0$, thus it suffices to show that $\frac{\partial \Phi}{\partial_+ w} < 0$. We have:

$$\frac{\partial \Phi}{\partial_+ w} = \lambda_0 F_p(c)\frac{\alpha}{\beta} t_w'(N_I) N_0 f_d(w) - \lambda_0[1 - F_p(c)]\frac{\alpha\lambda_0}{\beta\mu} t_w''(N_I) f_p(c) N_0 f_d(w) < 0.$$

This completes the proof. $\square$

E: Proof of Proposition 4

Proof. To solve (17), we first obtain the solution to (9) and denote it as $(\tilde{p}_f, \tilde{p}_d, \tilde{\lambda}, \tilde{N})$. If $\tilde{N} \leq N_{cap}$, then the cap constraint (18c) is inactive, and (17) reduces to (9).

If $\tilde{N} > N_{cap}$, then (18c) is active. In this case, (18a) and (18b) is equivalent to

$$\begin{cases} \lambda = \lambda_0 \left[1 - e_p \left(\dfrac{\alpha M}{\sqrt{N_{cap} - \lambda/\mu}} + \beta p_f \right) \right] & \text{(62a)} \\ N_{cap} = N_0 e_d \dfrac{\lambda p_d}{N_{cap}} & \text{(62b)} \end{cases}$$

According to (62b), $\lambda p_d = \dfrac{N_{cap}^2}{N_0 e_d}$. Therefore, the objective function is $R = \lambda p_f - \dfrac{N_{cap}^2}{N_0 e_d}$. Apply implicit function theorem on (62a), we can obtain $\dfrac{\partial \lambda}{\partial p_f}$ and further $\dfrac{\partial R}{\partial p_f}$. The first order condition becomes

$$\lambda_0 e_p \beta \mu p_f = \lambda_0 e_p \alpha M \frac{\lambda}{2(N_{cap} - \lambda/\mu)^{3/2}} + \lambda \mu. \tag{63}$$

Uniqueness: we first show that there is at most one set of $\lambda > 0$, $p_f > 0$ and $p_d > 0$ that satisfy (62a), (62b), and (63). Assume not, i.e., (λ, p_f, p_d) and (λ', p_f', p_d') are both solutions to (62a), (62b), and (63). If $p_f = p_f'$, then it is easy to verify that $p_d = p_d'$ and $\lambda' = \lambda$. Therefore, we consider $p_f' > p_f$ without loss of generality. Based on (62a), we have $\lambda' < \lambda$. However, (63) dictates that $\lambda' > \lambda$, A contradiction.

Existence and Optimality: Note that for any $p_f \geq 0$ such that

$$e_p \left(\frac{\alpha M}{\sqrt{N_{cap}}} + \beta p_f \right) < 1. \tag{64}$$

There is a unique $\lambda > 0$ that satisfies (62a). This is because the righthand side of (62a) is a concave function λ which has a unique intersection with the left-hand side of (62a). We can therefore view λ as a function of p_f determined by (62a), denoted by $\lambda = h_1(p_f)$. Apply implicit function theorem on (62a), we have $\dfrac{\partial \lambda}{\partial p_f} < 0$. Therefore, it is a decreasing function such that $h_1(0) > 0$, and it can be verified that $h_1(p_f) \to 0$ for sufficiently large p_f that satisfies (64). On the other hand, (63) prescribes λ as an increasing function of p_f. We denote it as as $\lambda = h_2(p_f)$, and we have $h_2(0) = 0$. To prove existence, it suffices to show that $h_1(p_f)$ intersects with $h_2(p_f)$ at some $p_f^* > 0$, which is clearly true.

p_f^* is either local minimum or local maximum. Note that p_f is bounded between 0 and an upper bound determined by (64). On the boundary the objective value R is smaller than that in the interior where $\lambda > 0$ and $p_f > 0$ (since the revenue λp_f on the boundary is 0). This indicates that p_f^* has to be local maximum. Since there is a unique solution to the first order conditions, p_f^* is the globally optimal solution. This completes the proof.

□

F: Proof of Lemma 1

Proof. Let (p_f^*, λ^*, N^*), $(\bar{p}_f, \bar{\lambda}, \bar{N})$ and $(\hat{p}_f, \hat{\lambda}, \hat{N})$ be the optimal solution (50), (51) and (53), respectively. Define R^*, $\bar{R}$ and $\hat{R}$ as the corresponding optimal values. Note that $R^* \leq \bar{R}$ and $R^* \leq \hat{R}$. This is because the objective value of (50) is smaller than that of (51) and (53). We consider the following three cases:

- $\Pi_1(\bar{N}) < \Pi_2(\bar{N})$. Given N, the optimization problems (51) and (53) are equivalent (since the second term of the objective functions are constants). Therefore, the corresponding optimal solution (i.e., λ and p_f) are the same when N are the same, and $\Pi_1(N) - \Pi_2(N) = wN - g(N)$ for any N. This indicates that $g(\bar{N}) > w\bar{N}$. It further implies that the objective value of (50) can attain $\bar{R}$ when $N = \bar{N}$. Since R^* is upper bounded by $\bar{R}$, we conclude that $N^* = \bar{N}$.

- $\Pi_2(\hat{N}) < \Pi_1(\hat{N})$. In this case we have $N^* = \hat{N}$. Proof is the same as case (a).

- $\Pi_1(\bar{N}) \geq \Pi_2(\bar{N})$ and $\Pi_2(\hat{N}) \geq \Pi_1(\hat{N})$. In this case, we can show that $g(N^*) = wN^*$. Assume not, then without loss of generality, consider the case where $g(N^*) > wN^*$. Consider the following problem:

$$\max_{p_f, N} \quad \lambda p_f - g(N) \tag{65}$$

$$\begin{cases} \lambda = \lambda_0 \left[1 - F_p \left(\dfrac{\alpha M}{\sqrt{N - \lambda/\mu}} + \beta p_f \right) \right] & \text{(66a)} \\ g(N) \geq wN & \text{(66b)} \end{cases}$$

We conclude that N^* is the optimal solution to (65). This is because if not, then there exists another solution that satisfies (66) and obtains a higher value than R^*, which contradicts with the fact that N^* is optimal solution to (50). Since $g(N^*) > wN^*$, the constraint (66b) is inactive at the optimal solution to (65), therefore, it is equivalent to (51), i.e., $N^* = \bar{N}$. In this case, $g(N^*) > wN^*$ implies $g(\bar{N}) > w\bar{N}$. Since $\Pi_1(N) - \Pi_2(N) = wN - g(N)$ for any N, we have $\Pi_1(\bar{N}) < \Pi_2(\bar{N})$. This contradicts with the assumption that $\Pi_1(\bar{N}) \geq \Pi_2(\bar{N})$. Therefore, $g(N^*) = wN^*$. This indicates that both (16b) and (16c) are active at the optimal solution to (15), which leads to (55).

This completes the proof.

□